POLYPEPTIDES INVOLVED IN THE BIOSYNTHESIS OF COBALAMINS AND/OR COBAMIDES, DNA SEQUENCES CODING FOR THESE POLYPEPTIDES, PREPARATION METHOD AND THEIR USE

The present invention relates to new polypeptides

5　involved in the biosynthesis of cobalamins and/or cobamides, and especially of coenzyme B_{12}. It also relates to the genetic material responsible for the expression of these polypeptides, as well as to a method by means of which they may be prepared. It relates, lastly, to a method

10　for amplification of the production of cobalamins, and more especially of coenzyme B_{12}, by recombinant DNA techniques.

Vitamin B_{12} belongs to the B group of vitamins. It is a water-soluble vitamin which has been identified as the factor enabling patients suffering from pernicious anaemia

15　to be treated. It is generally prescribed to stimulate haematopoiesis in fatigue subjects, but it is also used in many other cases comprising liver disorders and nervous deficiencies or as an appetite stimulant or an active principle with tonic activity, as well as in dermatology

20　(Berck, 1982, Fraser et al., 1983). In the industrial rearing of non-ruminant animals, the feed being essentially based on proteins of vegetable origin, it is necessary to incorporate vitamin B_{12} in the feed rations in amounts of 10 to 15 mg per tonne of feed (Barrère et al., 1981).

25　Vitamin B_{12} belongs to a class of molecules known as cobalamins, the structure of which is presented in Figure 1. Cobamides differ from cobalamins in the base of

the lower nucleotide, which is no longer 5,6-
dimethylbenzimidazole but another base, e.g.
5-hydroxybenzimidazole for vitamin B_{12}-factor III
synthesised, inter alia, by <u>Clostridium thermoaceticum</u> and
5   <u>Methanosarcina barkeri</u> (Iron et <u>al</u>., 1984). These
structural similarities explain the fact that the metabolic
pathways of biosynthesis of cobalamins and cobamides are,
for the most part, shared.

Cobalamins are synthesised almost exclusively by
10   bacteria, according to a complex and still poorly
understood process which may be divided into four steps
(Figure 2):

i) synthesis of uroporphyrinogen III (or uro'gen
III), then

15   ii) conversion of uro'gen III to cobyrinic acid,
followed by

iii) conversion of the latter to cobinamide, and

iv) construction of the lower nucleotide
loop with incorporation of the particular base
20   (5,6-dimethylbenzimidazole in the case of cobalamins).

For coenzyme B_{12}, it is probable that the addition
of the 5'-deoxyadenosyl group occurs shortly after the
corrin ring-system is synthesised (Huennekens et <u>al</u>.,
1982).

25   In the case of cobamides, only the step of
synthesis and incorporation of the lower base is different.

The first part of the biosynthesis of cobalamins
is very well known, since it is common to that of haemes as
well as to that of chlorophylls (Battersby et al., 1980).
It involves, successively, δ-aminolevulinate synthase (EC
5   2.3.137), δ-aminolevulinate dehydrase (EC 4.2.1.24),
porphobilinogen deaminase (EC 4.3.1.8) and uro'gen III
cosynthase (EC 4.2.1.75), which convert succinyl-CoA and
glycine to uro'gen III. However, the first step takes place
in some organisms [e.g. E. coli (Avissar et al., 1989) and
10   in methanogenic bacteria (Kannangara et al., 1989), for
example] by the conversion by means of a multi-enzyme
complex of glutamic acid to δ-aminolevulinic acid.

Between uro'gen III and cobyrinic acid, only
three intermediate derivatives have been purified to date;
15   they are the factors FI, FII and FIII, which are oxidation
products, respectively, of the three intermediates
precorrin-1, precorrin-2 and precorrin-3, which correspond
to the mono-, di- and trimethylated derivatives of uro'gen
III (Figure 3); these intermediates are obtained by
20   successive transfers of methyl groups from SAM (S-adenosyl-
L-methionine) to uro'gen III at positions C-2, C-7 and C-
20, respectively. The other reactions which take place to
give cobyrinic acid are, apart from five further transfers
of methyl groups from SAM at C-17, C-12, C-1, C-15 and C-5,
25   elimination of the carbon at C-20, decarboxylation at C-12
and insertion of a cobalt atom (Figure 4). These
biosynthetic steps have been deduced from experiments

performed _in vitro_ on acellular extracts of
Propionibacterium shermanii or of _Clostridium_
tetanomorphum. In these extracts, cobyrinic acid is
obtained by conversion of uro'gen III after incubation

5   under suitable anaerobic conditions (Batterby et _al_.,
1982). No intermediate between precorrin-3 and cobyrinic
acid capable of being converted to corrinoids by subsequent
incubation with extracts of cobalamin-producing bacteria
has been isolated to date. The difficulty of isolating and

10   identifying these intermediates is linked to

> i) their great instability,

> ii) their sensitivity to oxygen, and

> iii) their low level of accumulation _in vivo_.

In this part of the pathway, only one enzyme of _Pseudomonas_

15   _denitrificans_ has been purified and studied; it is
SAM:uro'gen III methyltransferase (Blanche et _al_., 1989),
referred to as SUMT.

> Between cobyrinic acid and cobinamide, the
following reactions are performed:

20   > i) addition of the 5'-deoxyadenosyl group (if
coenzyme B_{12} is the compound to be synthesised),

> ii) amidation of six of the seven carboxyl
functions by addition of amine groups, and

> iii) amidation of the last carboxyl function

25   (propionic acid chain of pyrrole ring D) by addition of
(R)-1-amino-2-propanol (Figure 2).

Whether there was really an order in the
amidations was not elucidated (Herbert et al., 1970).
Lastly, no assay of activity in this part of the pathway
has been described, except as regards the addition of the
5   5'-deoxyadenosyl group (Huennekens et al., 1982).

The final step of the biosynthesis of a
cobalamin, e.g. coenzyme B_{12}, comprises four successive
phases described in Figure 5 (Huennekens et al., 1982),
namely:

10   i) phosphorylation of the hydroxyl group of the
aminopropanol residue of cobinamide to cobinamide
phosphate, then

ii) addition of a guanosine diphosphate by
reaction with guanosine 5'-triphosphate; the compound
15   obtained is GDP-cobinamide (Friedmann, 1975), which

iii) reacts with 5,6-dimethylbenzimidazole,
itself synthesised from riboflavin, to give
adenosylcobalamin 5'-phosphate (Friedmann et al., 1968),
which

20   iv) on dephosphorylation leads to coenzyme B_{12}
(Schneider and Friedmann, 1972).

Among bacteria capable of producing cobalamins,
the following may be mentioned in particular:

Agrobacterium tumefaciens
25   Agrobacterium radiobacter

Bacillus megaterium

Clostridium sticklandii

Clostridium tetanomorphum

Clostridium thermoaceticum

Corynebacterium XG

Eubacterium limosum

5   Methanobacterium arbophilicum

Methanobacterium ivanovii

Methanobacterium ruminantium

Methanobacterium thermoautotrophicum

Methanosarcina barkeri

10   Propionobacterium shermanii

Protaminobacter ruber

Pseudomonas denitrificans

Pseudomonas putida

Rhizobium meliloti

15   Rhodopseudomonas sphaeroides

Salmonella typhimurium

Spirulina platensis

Streptomyces antibioticus

Streptomyces aureofaciens

20   Streptomyces griseus

Streptomyces olivaceus

At the industrial level, as a result of the great
complexity of the biosynthetic mechanisms, the production
of cobalamins, and especially of vitamin B_{12}, is exclusively
25   microbiological. It is carried out by large-volume cultures
of the bacteria Pseudomonas denitrificans,
Propionibacterium shermanii and Propionibacterium

freudenreichii (Florent, 1986). The strains used for the
industrial production are derived from wild-type strains;
they may have undergone a large number of cycles of random
mutation and then of selection of improved clones for the

5   production of cobalamins (Florent, 1986). The mutations are
obtained by mutagenesis with mutagenic agents or by
physical treatments such as treatments with ultraviolet
rays (Barrère et al., 1981). By this empirical method,
random mutations are obtained and improve the production of

10   cobalamins. For example, it is described that, from the
original strain of Pseudomonas denitrificans initially
isolated by Miller and Rosenblum (1960, US Patent
2,938,822), the production of this microorganism was
gradually increased in the space of ten years, by the

15   techniques mentioned above, from 0.6 mg/l to 60 mg/l
(Florent, 1986). For bacteria of the genus
Propionibacterium [Propionibacterium shermanii (ATCC
13673) and freudenreichii (ATCC 6207)], the same production
values appear to be described in the literature; e.g. a

20   production of 65 mg/l has been described (European Patent
87,920). However, no screen has yet been described enabling
either mutants overproductive of cobalamins or mutants
markedly improved in their production of cobalamins to be
readily selected or identified.

25   At the genetic level, little work has been
performed to date. The cloning of cob genes (coding for
enzymes involved in the biosynthetic process) has been

described in Bacillus megaterium (Brey et al., 1986).

Eleven complementation groups have been identified by

complementation of cob mutants of Bacillus megaterium with

plasmids carrying different fragments of Bacillus

megaterium DNA. These genes are grouped on the same locus,

carried by a 12-kb fragment.

Studies have also been carried out on the cob

genes of Salmonella typhimurium. Although the cloning of

these has not been described, it has been shown that almost

all the genes for cobalamin biosynthesis are grouped

together between minutes 40 and 42 of the chromosome (Jeter

and Roth, 1987). Only the cysG locus, which must permit the

conversion of uro'gen III to precorrin-2, does not form

part of this group of genes. However, the activity encoded

by this locus and also its biochemical properties have not

been described.

In addition, some phenotypes have been associated

with cob mutations. In Salmonella typhimurium and in

Bacillus megaterium, cob mutants no longer show growth on

minimum medium with ethanolamine as a carbon source or as a

nitrogen source (Roof and Roth, 1988). This is due to the

fact that an enzyme of ethanolamine catabolism,

ethanolamine ammonia-lyase (EC 4.3.1.7), has coenzyme B_{12} as

a cofactor; the cob mutants no longer synthesise coenzyme

B_{12}, and they can no longer grow with ethanolamine as a

carbon source and/or as a nitrogen source. metE mutants of

Salmonella typhimurium retain only a methylcobalamin-

dependent homocysteine methyltransferase (EC 2.1.1.13). cob
mutants of Salmonella typhimurium metE are auxotrophic for
methionine (Jeter et al., 1984).

In Pseudomonas denitrificans and Agrobacterium
tumefaciens, phenotypes associated with a total deficiency
of cobalamin synthesis have not been described to date.

Finally, work on Pseudomosas denitrificans
(Cameron et al., 1989) has led to the cloning of DNA
fragments carrying cob genes of this bacterium. These are
distributed in four complementation groups carried by at
least 30 kb of DNA. At least fourteen complementation
groups have been identified by heterologous complementation
of cob mutants of Agrobacterium tumefaciens and of
Pseudomonas putida with DNA fragments of Pseudomonas
denitrificans carrying cob genes.

However, hitherto, none of these genes has been
purified, and no nucleotide sequence has been described.
Similarly, no protein identification nor any catalytic
function attributed to the product of these genes has been
described. Furthermore, no improvement in production of
cobalamins by recombinant DNA techniques could be obtained.
The amplification of cob genes of Bacillus megaterium does
not bring about, in the strain from which they have been
cloned, an improvement in production of cobalamins (Brey et
al., 1986). In Salmonella typhimurium, physiological
studies have been carried out in order to determine
conditions under which a strong transcription of the cob

genes studied was observed (Escalante and Roth, 1987).
Under these conditions, there is no improvement in the
production of cobalamins, although genes of the
biosynthetic pathway are more expressed than under standard
5   culture conditions.

The present invention results from the precise
identification of DNA sequences coding for polypeptides
involved in the biosynthesis of cobalamins and/or
cobamides. A subject of the invention hence relates to the
10   DNA sequences coding for the polypeptides involved in the
biosynthesis of cobalamines and/or cobamides. More
especially, the subject of the invention is the cobA, cobB,
cobC, cobD, cobE, cobF, cobG, cobH, cobI, cobJ, cobK, cobL,
cobM, cobN, cobO, cobP, cobQ, cobS, cobT, cobU, cobV, cobW,
15   cobX and corA genes, any DNA sequence homologous with these
genes resulting from the degeneracy of the genetic code,
and also DNA sequences, of any origin (natural, synthetic,
recombinant), which hybridise and/or which display
significant homologies with these sequences or with
20   fragments of the latter, and which code for polypeptides
involved in the biosynthesis of cobalamins and/or
cobamides. The subject of the invention is also the genes
containing these DNA sequences.

The DNA sequences according to the present
25   invention were isolated from an industrial strain,
Pseudomonas denitrificans SC510, derived from strain MB580
(US Patent 3,018,225), by complementation of cob mutants of

A. tumefaciens and P. putida; and of Methanobacterium
ivanovii. The clones obtained could be analysed precisely,
in particular by mapping using insertions of a derivative
of transposon Tn5. These genetic studies have enabled the

5 cob or cor genes to be localised on the restriction map and
their sequencing to be carried out. An analysis of the open
reading frames then enabled the coding regions of these DNA
fragments to be demonstrated.

The subject of the present invention is also the

10 use of these nucleotide sequences for cloning the cob genes
of other bacteria. In effect, it is known that, for
proteins catalysing the same activities, sequences are
conserved, the divergence being the evolutionary divergence
(Wein-Hsiung et al., 1985). It is shown in the present

15 invention that there is a significant homology between the
nucleotide sequences of different microorganisms coding for
polypeptides involved in the biosynthesis of cobalamins
and/or cobamides. The differences which are seen result
from the evolutionary degeneracy, and from the degeneracy

20 of the genetic code which is linked to the percentage of GC
in the genome of the microorganism studied (Wein-Hsiung et
al., 1985).

According to the present invention, a probe may
be made with one or more DNA sequences of Pseudomonas

25 denitrificans in particular, or with fragments of these, or
with similar sequences displaying a specific degree of
degeneracy in respect of the use of the codons and the

percentage of GC in the DNA of the bacterium which it is

desired to study. Under these conditions, it is possible to

detect a specific hybridisation signal between the probe

and fragments of genomic DNA of the bacterium studied; this

5   specific hybridisation signal corresponds to the

hybridisation of the probe with the isofunctional cob genes

of the bacterium. The cob genes as well as their products

may then be isolated, purified and characterised. The

invention thus provides a means enabling access to be

10   gained, by hybridisation, to the nucleotide sequences and

the polypeptides involved in the biosynthesis of cobalamins

and/or cobamides of any microorganism.

The subject of the present invention is also a

recombinant DNA containing at least one DNA sequence coding

15   for a polypeptide involved in the biosynthesis of

cobalamins and/or cobamides, and in particular a

recombinant DNA in which the said sequence or sequences are

placed under the control of expression signals.

In this connection, promoter regions may, in

20   particular, be positioned at the 5' end of the DNA

sequence. Such regions may be homologous or heterologous to

the DNA sequence. In particular, strong bacterial promoters

such as the promoter of the tryptophan operon Ptrp or of

the lactose operon Plac of E. coli, the leftward or

25   rightward promoter of bacteriophage lambda, the strong

promoters of phages of bacteria such as Corynebacteria, the

functional promoters in Gram-negative bacteria such as the

Ptac promoter of E. coli, the PxylS promoter of the xylene catabolism genes of the TOL plasmid and the amylase promoter of Bacillus subtilis Pamy may be used. Promoters derived from glycolytic genes of yeasts may also be mentioned, such as the promoters of the genes coding for phosphoglycerate kinase, glyceraldehyde-3-phosphate dehydrogenase, lactase or enolase, which may be used when the recombinant DNA is to be introduced into a eukaryotic host. A ribosome binding site will also be positioned at the 5' end of the DNA sequence, and it may be homologous or heterologous, such as the ribosome binding site of the cII gene of bacteriophage lambda.

Signals necessary to transcription termination may be placed at the 3' end of the DNA sequence.

The recombinant DNA according to the present invention may then be introduced directly into a host cell compatible with the chosen expression signals, or be cloned into a plasmid vector to enable the DNA sequence in question to be introduced in a stable manner into the host cell.

Another subject of the invention relates to the plasmids thereby obtained, containing a DNA sequence coding for a polypeptide involved in the biosynthesis of cobalamins and/or cobamides. More specifically, these plasmids also contain a functional replication system and a selectable marker.

The subject of the invention is also the host cells into which one or more DNA sequences as defined above, or a plasmid as defined hereinbefore, has/have been introduced.

5   Another subject of the invention relates to a method for production of polypeptides involved in the biosynthesis of cobalamins and/or cobamides. According to this method, a host cell is transformed with a DNA sequence as described above, this transformed cell is cultured under 10   conditions for expression of the said sequence and the polypeptides produced are then recovered.

The host cells which may be used for this purpose are either prokaryotes or eukaryotes, animal cells or plant cells. Preferably, they will be chosen from bacteria, and 15   especially bacteria of the genus E. coli, P. denitrificans, A. tumefaciens or R. meliloti.

Another use of the DNA sequences according to the present invention lies in a method for amplification of the production of cobalamins and/or cobamides, by recombinant 20   DNA techniques. In effect, if the limitation of the metabolic flux of the biosynthesis of cobalamins and/or cobamides is due to a limitation in the activity of an enzyme in the biosynthetic pathway, an increase in this activity by increasing the expression of this same enzyme 25   using recombinant DNA techniques (gene amplification, substitution of the transcription/translation signals with more effective signals, etc.) will lead to an increase in

the biosynthesis of cobalamins and/or cobamides. It is also possible that the limitation of the production of cobalamins and/or cobamides results from a biochemical regulation. In this case, the cob gene or genes
5   corresponding to the regulated enzyme may be specifically mutagenised in vitro in order to obtain mutated genes whose products will have lost the regulation mechanisms impeding an improvement in the production.

The method according to the present invention
10 consists in transforming a microorganism productive of cobalamins and/or cobamides, or only potentially productive of these compounds (i.e. deficient in one or more steps of the biosynthesis), with a DNA sequence as defined above, then in culturing this microorganism under conditions for
15 expression of the said sequence and for synthesis of cobalamins and/or cobamides, and lastly in recovering the cobalamins and/or cobamides produced. Such a method is applicable, in particular, to all the productive microorganisms mentioned on pages 5 and 6, and more
20 specifically to microorganisms of the genus P. denitrificans, Rhizobium meliloti, or Agrobacterium tumefaciens. In a preferred embodiment, the microorganism is P. denitrificans, and especially strain SC510. As regards potentially productive microorganisms, the DNA
25 sequences used will be those corresponding to the steps of the biosynthesis which the microorganism cannot carry out.

Using the present invention, and by the various strategies described above, an improvement in the production of cobalamins and/or cobamides may be obtained for any microorganism productive or potentially productive

5   of cobalamins and/or cobamides. It will suffice to culture this recombinant microorganism under suitable conditions for the production of cobalamins and for the expression of the DNA sequences introduced. This culturing may be carried out batchwise or alternatively in continuous fashion, and

10   the purification of the cobalamins may be carried out by the methods already used industrially (Florent, 1986). These methods comprise, inter alia:

        i) solubilisation of the cobalamins and their conversion to their cyano form (e.g. by heat treatment of

15   the fermentation must, with potassium cyanide in the presence of sodium nitrite), then

        ii) purification of the cyanocobalamins in various steps which can be, e.g.

        a)   adsorption on different substrates such as Amberlite

20   IRC-50, Dowex 1 x 2 or Amberlite XAD-2, followed by an elution with a water/alcohol or water/phenol mixture, then

        b)   extraction in an organic solvent, and lastly

        c)   precipitation or crystallisation from the organic phase, either by the addition of reagents or dilution in a

25   suitable solvent, or by evaporation.

        The present invention shows, furthermore, that it is possible by recombinant DNA techniques to improve the

cobalamin production of a bacterium productive of cobalamins by cumulating improvements. This amounts to obtaining a first improvement as described above, and then in improving this improvement, still using recombinant DNA techniques, i.e., e.g. by amplifying genes for cobalamin biosynthesis.

Another subject of the present invention relates to the polypeptides involved in the biosynthesis of cobalamins and/or cobamides. In particular, the subject of the present invention is all polypeptides, or derivatives or fragments of these polypeptides, which are encoded by the DNA sequences described above, and which are involved in the pathway of biosynthesis of cobalamins and/or cobamides. The amino acid sequence of these polypeptides is described, as well as some of their physicochemical properties. An enzymatic activity or specific properties have also been associated with each of them.

In this connection, the subject of the invention is the polypeptides participating in the conversion of precorrin-3 to cobyrinic acid a,c-diamide, and more especially in the transfer of a methyl group from SAM to positions C-1, C-5, C-11, C-15 and C-17.

The subject of the invention is also the polypeptides:

· participating in the conversion of cobyric acid to cobinamide, or

· possessing an S-adenosyl-L-methionine:precorrin-2 methyltransferase (SP2MT) activity, or

· possessing a cobyrinic and/or hydrogenobyrinic acid a,c-diamide synthase activity, or

· possessing a precorrin-8x mutase activity, or

· possessing a nicotinate-nucleotide: dimethylbenzimidazole phosphoribosyltransferase activity, or

· possessing a cobalamin-5'-phosphate synthase activity, or

· possessing a cobyric acid synthase activity, or

· possessing a cob(I)alamin adenosyl-transferase activity, or

· possessing a precorrin-6x reductase activity, or

· participating in the conversion of hydrogenobyrinic acid a,c-diamide to cobyrinic acid a,c-diamide.

Advantageously, the subject of the invention is a polypeptide chosen from the COBA, COBB, COBC, COBD, COBE, COBF, COBG, COBH, COBI, COBJ, COBK, COBL, COBM, COBN, COBO, COBP, COBQ, COBS, COBT, COBU, COBV, COBW, COBX and CORA proteins presented in Figures 15, 16, 40, 41 and 47.

Furthermore, the use of the hybridisation probes described above makes it possible, from genes isolated in other microorganisms, to characterise and isolate the

isofunctional polypeptides of other microorganisms. In this manner, the present invention shows that the sequence of a COB protein of <u>Pseudomonas denitrificans</u> is significantly homologous with the protein sequences of other

5   microorganisms displaying the same type of activity. Between these COB proteins catalysing the same reaction in different microorganisms, only the evolutionary distances have introduced variations (Wein-Hsiung et <u>al</u>., 1985). The subject of the present invention is also these

10   isofunctional polypeptides.

The assignment of a particular enzymatic activity is the result of an analysis which may be performed according to various strategies. In particular, <u>in vitro</u> affinity studies with respect to SAM (S-adenosyl-L-

15   methionine) make it possible to assign a methyl transferase activity to a protein capable of binding SAM, and hence to assign its involvement in one of the steps of transfer of methyl groups which occur between uro'gen III and cobyrinic acid. Another means of assessing the activity of these

20   polypeptides consists in assaying the intermediates in the pathway of biosynthesis of cobalamins which are accumulated in mutants incapable of expressing these polypeptides (identified by complementation experiments). These analyses enable it to be deduced that the polypeptide in question

25   has the accumulated intermediate as its substrate, thereby enabling its activity in the biosynthetic pathway to be situated and defined. The present invention also describes

a method for assaying the enzymatic activities of the
biosynthetic pathway, applicable to any strain productive
of cobalamins and/or cobamides. These assays enable the
enzymatic activity assayed to be purified from any strain

5   productive of these compounds. From this purified activity,
the NH_2-terminal sequence of the COB protein in question,
or alternatively that of the subunits of this protein, may
be determined, thereby enabling the structural gene or
genes which code for the activity in question to be

10   identified. For _Pseudomonas denitrificans_, the structural
genes which code for activities of the biosynthetic pathway
are identified by finding, for each NH_2-terminal sequence,
the COB protein having the same NH_2-terminal sequence.

The present invention also describes a method

15   enabling intermediates in the pathway of biosynthesis of
cobalamins or of other corrinoids to be identified and
assayed in strains productive of cobalamins. These
intermediates may be assayed both in culture musts and in
the cells themselves. The intermediates which may be

20   assayed are all the corrinoids which occur in the
biosynthetic pathway after cobyrinic acid, namely, apart
from cobyrinic acid, cobyrinic acid monoamide, cobyrinic
acid diamide, cobyrinic acid triamide, cobyrinic acid
tetraamide, cobyrinic acid pentaamide, cobyric acid,

25   cobinamide, cobinamide phosphate, GDP-cobinamide, coenzyme
B_{12} phosphate and coenzyme B_{12}. The non-adenosylated forms of
these products may also be assayed by this technique.

Other subjects and advantages of the present invention will become apparent on reading the examples and the drawings which follow, which are to be considered as illustrative and not limiting.

5

Definition of the terms employed and abbreviations.

ATP: adenosine 5'-triphosphate

bp: base pairs

BSA: bovine serum albumin

10 CADAS: cobyrinic acid a,c-diamide synthase

cluster: group of genes

Cob: corresponds to the phenotype with a reduced level (at least 10-fold lower than the control) of production of cobalamins

15 cob gene: gene involved in the biosynthesis of cobalamins and/or cobamides from uro'gen III

COB protein: protein participating either as a catalyst in the pathway of biosynthesis of cobalamins, or as a regulatory protein in

20 the network of regulation of the cob genes, or both.

cor gene: gene involved in the biosynthesis of corrinoids from uro'gen III

COR protein: protein participating either as a catalyst

25 in the pathway of biosynthesis of corrinoids, or as a regulatory protein in

the network of regulation of the <u>cor</u> genes, or both

Corrinoids: cobyrinic acid derivatives possessing the corrin ring-system

dGTP: 2'-deoxyguanosine 5'-triphosphate

DMBI: dimethylbenzimidazole

dNTP: 2'-deoxyribonucleoside 5'-triphosphates

DTT: dithiothreitol

HPLC: high performance liquid chromatography

kb: kilobases

NN:DMBI PRT: nicotinate-nucleotide: dimethylbenzimidazole phosphoribosyltransferase

ORF: open reading frame

recombinant DNA: set of techniques making it possible either to combine within the same microorganism DNA sequences which are not naturally so combined, or to mutagenise specifically a DNA fragment

SAM: S-adenosyl-L-methionine

SDS: sodium dodecyl sulphate

SP_2MT: SAM-L-methionine:precorrin-2 methyltransferase

Stop codon: translation termination codon

SUMT: SAM:uro'gen III methyltransferase

Uro'gen III: uroporphyrinogen III

Legends to the figures:

Figure 1: Structure of coenzyme B_{12}; the 5'-deoxyadenosyl group is replaced by a CH_3 group for

methylcobalamin, by a cyano group for
cyanocobalamin, by a hydroxyl group for
hydroxocobalamin.

Figure 2: Biosynthesis of cobalamins and various steps
of this biosynthesis. X: axial ligands of the
cobalt; the ligand at a may be different from
the ligand at b. R: ligand at a of the cobalt
which defines the cobalamin type (see Figure
1).

Figure 3: Structures of uro'gen III, precorrin-1,
precorrin-2 and precorrin-3.

Figure 4: Structural formulae of uro'gen III and
cobyrinic acid. Between uro'gen III and
cobyrinic acid, there occur 8 SAM-dependent
methyl transfers successively at C-2, C-7,
C-20, C-17, C-12, C-1, C-15 and C-5, a
decarboxylation at C-12, elimination of the
carbon at C-20 and insertion of the cobalt
atom. X: axial ligands of the cobalt; the
ligand at a may be different from the ligand
at b.

Figure 5: Final steps of the biosynthesis of cobalamins.
In order to clarify the diagram, details of
the corrin ring-system have been omitted. The
five enzymatic steps are represented: 1,
cobinamide kinase; 2, cobinamidephosphate
guanylyltransferase; 3, cobalamin-5'-phosphate

synthase; 4, cobalamin-5'-phosphate

phosphohydrolase; 5, nicotinatenucleotide:DMBI

phosphoribosyltransferase.

Figure 6:   Restriction maps of the 5.4-kb ClaI-HindIII-

HindIII-HindIII, 8.7-kb EcoRI, 4748-bp SalI-

SalI-SalI-SalI-SalI-BglI and 3855-bp SstI-

SstI-BamHI fragments. Only the 20 restriction

enzymes which cut the DNA least frequently are

shown. The cleavage sites of each enzyme are

indicated by a vertical line.

Figure 7:   Nucleotide sequence of both strands of the

5378-bp ClaI-HindIII-HindIII-HindIII fragment

of Pseudomonas denitrificans. The strand

situated at the top is to be read from 5' to

3' in the left-to-right direction which

corresponds to the left-to-right orientation

of the sequenced fragment presented in Figure

6. The ClaI site occurs at position 23

(beginning of the cleavage site) since, in

this sequence, there occur PstI, SalI and XbaI

restriction sites which have appeared during

clonings in multisites with a view to

sequencing. The sequence of the ClaI-HindIII-

HindIII-HindIII fragment hence begins at

position 23.

Figure 8:   Nucleotide sequence of both strands of the

8753-bp EcoRI fragment of Pseudomonas

denitrificans. The strand situated at the top
is to be read from 5' to 3' in the left-to-
right direction which corresponds to the left-
to-right orientation of the fragment of the

5   restriction map presented in Figure 6.

Figure 9:   Analysis of the probabilities of the coding
frames on the basis of codon preference using
the programme of Staden and MacLachlan (1982)
on the 6 reading frames of the 5378-bp ClaI-

10   HindIII-HindIII-HindIII fragment. For the
frames belonging to the same coding strand,
the most probable frame corresponds to that in
which a dotted line, not interrupted by stop
codons, is placed under the probability line

15   for this frame.

1. Sequence extending from nucleotide 1 to
nucleotide 1200. By means of this analysis,
open reading frame 1 is identified. It begins
at the ATG at position 549 and ends at the TGA

20   at position 1011.

2. Sequence extending from nucleotide 1000 to
nucleotide 2200. By means of this analysis,
open reading frame 2 is identified. It begins
at the ATG at position 1141 and ends at the

25   TGA at position .1981.

3. Sequence extending from nucleotide 1800 to
nucleotide 3400. By means of this analysis,

open reading frame 3 is identified. It begins at the ATG at position 1980 and ends at the TGA at position 3282.

4. Sequence extending from nucleotide 3000 to nucleotide 4500. By means of this analysis, open reading frame 4 is identified. It begins at the ATG at position 3281 and ends at the TGA at position 4280.

5. Sequence extending from nucleotide 3800 to nucleotide 5378. By means of this analysis, open reading frame 5 is identified. It begins at the GTG at position 4284 and ends at the TGA at position 5253.

Figure 10: Analysis of the probabilities of the coding frames on the basis of codon preference using the programme of Staden and MacLachlan (1982) on the 6 reading frames of the 8753-bp EcoRI fragment. For the frames belonging to the same coding strand, the most probable frame corresponds to that in which a dotted line, not interrupted by stop codons, is placed under the probability line for this frame.

1. Sequence extending from nucleotide 650 to nucleotide 1650. By means of this analysis, open reading frame 6 is identified. It begins

at the ATG at position 736 and ends at the TGA at position 1519.

2. Sequence extending from nucleotide 1400 to nucleotide 3100. By means of this analysis, open reading frame 7 is identified. It begins at the ATG at position 1620 and ends at the TAG at position 2997.

3. Sequence extending from nucleotide 2700 to nucleotide 3700. By means of this analysis, open reading frame 8 is identified. It begins at the ATG at position 3002 and ends at the TGA at position 3632.

4. Sequence extending from nucleotide 3500 to nucleotide 4100. By means of this analysis, open reading frame 9 is identified. It begins at the GTG at position 3631 and ends at the TGA at position 4366.

5. Sequence extending from nucleotide 4150 to nucleotide 5150. By means of this analysis, open reading frame 10 is identified. It begins at the ATG at position 4365 and ends at the TGA at position 5127.

6. Sequence extending from nucleotide 5000 to nucleotide 6000. By means of this analysis, open reading frame 11 is identified. It begins at the ATG at position 5893 and ends at the TAG at position 5110.

7. Sequence extending from nucleotide 5700 to nucleotide 7200. By means of this analysis, frame 12 is identified. It begins at the ATG at position 5862 and ends at the TAA at position 7101.

8. Sequence extending from nucleotide 7000 to nucleotide 8000. By means of this analysis, open reading frame 13 is identified. It begins at the ATG at position 7172 and ends at the TTG at position 7931.

Figure 11: Construction of plasmids pXL556, pXL545 and pXL723.

A 2.4-kb ClaI-EcoRV fragment containing the cobA and cobE genes is excised from the 5.4-kb fragment and then purified. An EcoRI linker is added at the EcoRV site and the fragment is then inserted into pXL59 between the ClaI-EcoRI sites. The plasmid thereby constructed is designated pXL556.

The construction is comparable for pXL545: a 1.9-kb ClaI-HindIII-HindIII fragment is excised from the 5.4-kb fragment and then purified. This fragment contains only the CobE gene. An EcoRI linker is added at the HindIII site and the fragment is then inserted into pXL59 between the ClaI-EcoRI sites.

pXL723 is constructed as follows: a 2.3-kb EcoRI-HindIII fragment is excised from the 5.4-kb fragment and purified, and the ends are then filled in with the large fragment of E. coli DNA polymerase I. This fragment is cloned into pRK290 (Ditta et al., 1981) digested with EcoRI and then treated with the large fragment of E. coli DNA polymerase I in order to fill in the ends.

The restriction sites which are shown in brackets correspond to sites which have disappeared after treatment with the large fragment of E. coli DNA polymerase I.

1, PstI-SstI fragment of RSF1010 (De Graff et al., 1978); 2, PstI-BamHI fragment of pACYC177 (Bagdasarian et al., 1981); 3, BamHI-SstI fragment containing the lactose operon of E. coli without its promoter, the operator, the translation initiation site and the first 8 non-essential codons of lacZ (Casadaban et al., 1983); 4, Sau3AI fragment of Pseudomonas putida KT2440 (Bagdasarian et al., 1981); ori, origin of replication; nic, relaxation site; mob, locus essential for mobilisation; Km', kanamycin resistance gene (Bagdasarian et al., 1981); B, BamHI; C, ClaI; E, EcoRI; H,

HindIII; P, PstI; S, SstI; Sa, SalI; X, XhoI;

Xb, XbaI.

Figure 12:   Studies of the insertions of transposons

Tn5Sp' and Tn5 into the 5378-bp fragment. The

insertions of transposon Tn5 into plasmid

pXL723 are shown as in Figure 14; those of

transposon Tn5Sp', into the chromosome of

strain G2 Rif', are boxed; the insertions into

the chromosome of SC510 Rif' of cassettes

carrying the kanamycin resistance gene (1630

and 1631) are shown with an arrow, according

to the orientation of transcription of the

kanamycin resistance gene, under the insertion

number. The open reading frames deduced from

the sequence are given in this figure (from

cobA to cobE); + or - signs are shown under

each insertion of transposon or of resistance

cassette to indicate that the insertion is

inactivating (-) or otherwise (+), i.e. for

the complementation of different mutants (the

case with the insertions of transposons Tn5),

or that the insertion abolishes the cobalamin

production of the strain in which it takes

place. There is an absence of complementation

when the recombinant mutant synthesises less

than threefold less cobalamins than the level

of synthesis of the strain from which the

mutant is derived. The inserts of plasmids pXL545, pXL1500, pXL1397 and pXL302 are shown with the restriction sites occurring at their ends. These inserts are cloned into broad host range plasmids, pXL435 and pXL59 (Cameron et al., 1989):

plasmid pXL545 corresponds to plasmid pXL545 described in Figure 11 with, in addition, the 2-kb BamHI fragment of pHP45 (Prentki and Krisch) containing a spectinomycin resistance gene cloned at the BamHI site of pXL545;

plasmid pXL1500 corresponds to the 4.2-kb BglII-SstI fragment presented in this figure, cloned at the BamHI and SstI sites of pKT230 (Bagdasarian et al., 1981); presented in Figure 30;

plasmid pXL1397 corresponds to the 2.4-kb HindIII-SstI fragment indicated in the figure, inserted between the HindIII and SstI sites of the multisite of pXL435 (Cameron et al., 1989) described in Figure 30; plasmid pXL302 corresponds to the 2.3-kb EcoRI-HindIII fragment as described in the figure, inserted between the EcoRI and HindIII sites of pXL59 (Cameron et al., 1989) described in Figure 30,

the HindIII site used being the site occurring in the cloning multisite of pXL59;

pXL723 is described in Figure 11, like pXL545.

+ or - signs are shown above each of these inserts to indicate whether there is complementation by the plasmid in question of the chromosomal insertions shown underneath. C, ClaI; E, EcoRI; H, HindIII; RV, EcoRV; Sau, Sau3AI; S, SstI.

Figure 13:   Construction of plasmids pXL253 and pXL367. The 8.7-kb EcoRI fragment is excised and then purified from plasmid pXL151. It is cloned at the EcoRI site of pKT230 to give pXL253. This same fragment is inserted at the EcoRI site of pRK290 (Ditta et al., 1981) to give pXL367. 1, PstI-SstI fragment of RSF1010 (De Graff et al., 1978); 2, PstI-BamHI fragment of pACYC177 (Bagdasarian et al., 1981); ori, origin of replication; nic, relaxation site; mob, locus essential for mobilisation (Bagdasarian et al., 1981); B, BamHI; C, ClaI; E, EcoRI; H, HindIII; P, PstI; S, SstI; Sa, SalI; X, XhoI; Xb, XbaI; tet', tetracycline resistance gene; Km', kanamycin resistance gene.

Figure 14:   Studies of the insertions of transposons Tn3lacZ and Tn5 into the 8.7-kb EcoRI fragment

cloned into pRK290 (Ditta et al., 1980). The

insertions of transposons Tn3lacZ are

underlined, in contrast to those of

transposons Tn5. The open reading frames

5   deduced from the sequence (cobF to cobM) are

given in this figure, and the eight groups of

inactivating insertions (numbered from 1 to 8)

are presented; + or - signs are shown under

each transposon insertion to indicate that the

10   insertion is inactivating (-) or otherwise (+)

for the complementation of different mutants.

There is an absence of complementation when

the recombinant mutant synthesises less than

threefold less cobalamins than the level of

15   synthesis of the strain from which the mutant

is derived. These groups of inactivating

insertions correspond to the following

mutants: 1, G615; 2, G614 and G616; 3, G613

and G614; 4, G620; 5, G638; 6, G610 and G609;

20   7, G612; 8, G611. These mutants are Cob

mutants of Agrobacterium tumefaciens already

described (Cameron et al., 1989). A

restriction map of the 8.7-kb fragment is

given at the bottom of the figure.

25   Figure 15:   The coding sequences of each of the genes of

the 5.4-kb fragment, cobA to cobE,

respectively, are indicated. The sequences of

the proteins COBA to COBE encoded by these sequences appear under their respective coding sequence, cobA to cobE. The amino acid composition of each protein, in number and in

5   percentage, respectively, of COBA to COBE, is presented, as well as the molecular weight, the index of polarity, the isoelectric point and the optical density at 260 nm and 280 nm of a solution containing 1 mg/ml of purified

10   protein. The hydrophilicity profile of each COBA to COBE protein, respectively, is shown; it was calculated on the basis of the programme of Hopp and Woods (1981). Positive values correspond to regions of the protein

15   which are hydrophilic. The position of the amino acids is indicated as abscissa, while the value of the index of hydrophilicity is shown as ordinate; when this value is positive, this indicates that the region of

20   the protein is hydrophilic.

Figure 16:   The coding sequences of each of the genes of the 8.7-kb fragment, cobF to cobM respectively, are indicated. The sequences of the COBF to COBM proteins encoded by these

25   sequences appear under their sequence. The legend is identical to that for Figure 15. NB. We have shown the COBF protein as beginning at

the ATG located at position 736; it is possible that the ATG located at position 751 is the true initiation codon of this protein.

Figure 17: Reaction catalysed by cobyrinic acid a,c-diamide synthase. CADAS catalyses the amidation of the carboxylic acid functions of the peripheral acetate chains a and c of cobyrinic acid to give cobyrinic acid diamide; the donor of the amine group used in the enzymatic test is L-glutamine; it gives L-glutamic acid on deamination. X corresponds to the axial ligands of the cobalt, which may be different from one another.

Figure 18: Reaction catalysed by SP_2MT. SP_2MT catalyses the transfer of a methyl from SAM to dihydrosirohydrochlorin or precorrin-2 to give precorrin-3. The methyl group is transferred to position C-20 of the porphyrin ring-system.

Figure 19: Structure of hydrogenobyrinic acid and of hydrogenobyrinic acid a,c-diamide.

Figure 20: Affinities of the COBA and COBF proteins for SAM. The curves give in arbitrary units the radioactivity at emergence from the TSK-125 column for each protein applied to this column. The retention times are indicated in minutes and the radioactivity peak

corresponding to free SAM is observed at the
time of 10 min 30 sec.

Figure 21:   Comparison of the sequences of COBA and COBI.
Only the regions 1, 2 and 3, of strong
homology, are presented. = signs are placed
between identical residues and - signs between
homologous residues (H K R, L I V M, A G S T, Y
F W, D E Q N B Z, P, C).

Figure 22:   Comparison of the primary sequences of the
proteins COBA of Pseudomonas denitrificans and
CYSG of E. coli. The alignment has been done
according to the programme of Kanehisa, 1984.
= signs are placed between identical residues
and - signs between homologous residues (H K R,
L I V M, A G S T, Y F W, D E Q N B Z, P, C). The
regions 1, 2 and 3 correspond to zones of
strong homology between the proteins.

Figure 23:   Comparison of the sequences of CYSG of E. coli
with COB proteins of Pseudomonas denitrificans
(COBA, COBF, COBI, COBJ, COBL and COBM). The
comparisons concern the regions 1, 2 and 3, of
strong homologies, which exist between CYSG,
COBA and COBI. The positions in the protein
sequences of the regions displaying homologies
are presented in the figure. We have
considered the following groups of homologous
residues: H K R, L I V M, A G S T, Y F W, D E

Q N B Z, P, C. If there are at least 3
homologous residues at the same position, we
have boxed these amino acids.

Figure 24: Construction of plasmids pXL1148 and pXL1149.

5   pXL1148 is constructed as follows: the 1.9-kb
BamHI-BamHI-SstI-SstI fragment of the 8.7-kb
fragment containing the cobH and cobI genes is
purified, and XbaI and EcoRI linkers are
placed at the BamHI and SstI ends

10   respectively. This fragment is then inserted
between the XbaI and EcoRI sites of the broad
host range plasmid pXL59 (Cameron et al.,
1989) to give plasmid pXL1148.

pXL1149 is constructed like pXL1148, apart

15   from the fact that the fragment initially
purified is the 1.5-kb BamHI-BamHI-SstI
fragment instead of the fragment additionally
containing the small 400-bp SstI fragment used
for pXL1148. The fragment then undergoes the

20   same enzymatic treatments and the same cloning
into pXL59.

1, PstI-SstI fragment of RSF1010 (De Graff et
al., 1978); 2, PstI-BamHI fragment of pACYC177
(Bagdasarian et al., 1981); 3, BamHI-SstI

25   fragment containing the lactose operon of E.
coli without promoter, operator, translation
initiation site and the first 8 non-essential

codons of _lacZ_ (Casadaban et _al._, 1983); 4,
_Sau_3AI fragment of _Pseudomonas putida_ KT2440
(Bagdasarian et _al._, 1981); _ori_, origin of
replication; _nic_, relaxation site; Km',

5   kanamycin resistance gene; _mob_, locus
essential for mobilisation (Bagdasarian et
_al._, 1981); B, _Bam_HI; C, _Cla_I; E, _Eco_RI; H,
_Hind_III; P, _Pst_I; S, _Sst_I; Sa, _Sal_I; X, _Xho_I;
Xb, _Xba_I.

10   Figure 25:   Total proteins of strains SC510 Rif', SC510
Rif' pKT230, SC510 Rif' pXL1148, SC510 Rif'
pXL1149 analysed in 10 % SDS-PAGE as
described. The bacteria were cultured for 4
days in PS4 medium, and lysates of the total

15   proteins were then made. Lane 1, SC510 Rif';
lane 2, SC510 Rif' pXL1149; lane 3, SC510 Rif'
pXL1148; lane 4, SC510 Rif' pKT230. The
molecular masses of the molecular mass markers
are indicated. The positions to which the COBI

20   and COBH proteins migrate are indicated.

Figure 26:   Construction of plasmids pXL1496 and pXL1546.
Plasmid pXL1496 enables the COBF protein to be
overexpressed in _E. coli_, and plasmid pXL1546
enables COBF to be overexpressed in

25   _Pseudomonas denitrificans_.
The 2.2-kb _Eco_RI-_Xho_I fragment is excised and
purified from the 8.7-kb fragment. It is

cloned at the EcoRI site of phage M13mp19 to give plasmid pXL1405. An NdeI site is then introduced by directed mutagenesis, as described above, at position 733 of this fragment; in this manner, an NdeI site occurs exactly on the presumed initiation codon of the cobF gene. The new plasmid thereby obtained is designated pXL1406. A 1.5-kb NdeI-SphI-SphI fragment, containing the cobF gene starting from its presumed initiation codon, is purified after partial digestion with the appropriate enzymes and ligated with the appropriate fragments of plasmid pXL694 (120-bp EcoRI-NdeI fragment containing expression signals of E. coli - see text - and 3.1-kb EcoRI-SphI fragment containing the ampicillin resistance gene, the replication functions of the plasmid and also the terminators of the rrnB operon of E. coli, as described in the text). The plasmid thereby constructed is designated pXL1496.

pXL1546 is constructed as follows: the 2-kb EcoRI-BamHI-BamHI fragment of pXL1496 is purified by partial digestion with the appropriate enzymes; this fragment contains the expression signals of E. coli, followed by the cobF gene and then the 5' portion of the

cobG gene, this portion itself being followed
by terminators of the rrnB operon of E. coli,
as described in the text. This fragment is
cloned into the multihost plasmid pKT230

5   (Bagdasarian et al., 1981) described in Figure
30. B, BamHI; C, ClaI; E, EcoRI; H, HindIII;
P, PstI; S, SstI, Sa, SalI; X, XhoI; Xb, XbaI;
Km', kanamycin resistance gene; Amp,
ampicillin resistance gene.

10   Figure 27:   Total proteins of strains SC510 Rif', SC510
Rif' pKT230, SC510 Rif' pXL1546 analysed in
10 % SDS-PAGE as described. The bacteria were
cultured for 4 days in PS4 medium, and lysates
of the total proteins were then made. Lane 1,

15   SC510 Rif'; lane 2, SC510 Rif' pKT230; lane 3,
SC510 Rif' pXL1546. The molecular masses of
the molecular mass markers are indicated. The
position to which the COBF protein migrates is
indicated.

20   Figure 28:   Total proteins of the strains E. coli B and E.
coli B pXL1496 analysed in 10 % SDS-PAGE as
described. Lane 1, E. coli pXL1496 cultured in
the absence of tryptophan; lane 2, E. coli
pXL1496 cultured under the same conditions in

25   the presence of tryptophan; lane 3, E. coli
cultured in the absence of tryptophan; lane 4,
E. coli cultured under the same conditions in

the presence of tryptophan. The molecular masses of the markers are indicated. The position of migration of the COBF protein is indicated.

5 Figure 29: Construction of plasmids pXL525 and pXL368. Plasmid pXL368 is constructed as follows: the 2.4-kb EcoRV-ClaI fragment (containing the cobA and cobE genes) is purified from plasmid pXL556, thereby enabling this fragment to be

10 obtained with a BamHI site and an XbaI site at the ends; this fragment is cloned into pXL203 at the BamHI and XbaI sites.

For the construction of pXL525, an XbaI linker is added at the EcoRI site situated at the

15 right-hand end of the 8.7-kb EcoRI fragment; this 8.7-kb EcoRI-XbaI fragment is then cocloned with the 2.4-kb EcoRI-XbaI fragment originating from pXL556 and containing cobA and cobE.

20 The restriction sites which are shown in brackets correspond to sites which have disappeared after treatment with the large fragment of E. coli DNA polymerase I.

1, PstI-SstI fragment of RSF1010 (De Graff et

25 al., 1978); 2, PstI-BamHI fragment of pACYC177 (Bagdasarian et al., 1981); ori, origin of replication; nic, relaxation site; mob, locus

essential for mobilisation; Km', kanamycin

resistance gene (Bagdasarian et al., 1981); B,

BamHI; C, ClaI; E, EcoRI; H, HindIII; P, PstI;

S, SstI; Sa, SalI; X, XhoI; Xb, XbaI; tet,

5   tetracycline resistance gene; Amp' and Amp,

ampicillin resistance gene.

Figure 30: Plasmids of the incompatibility group Q having

a broad host range in Gram-negative bacteria.

These plasmids are described in a previous

10   publication (Cameron et al., 1989) and are

used in the present invention.

1, PstI-SstI fragment of RSF1010 (De Graff et

al., 1978); 2, PstI-BamHI fragment of pACYC177

(Bagdasarian et al., 1981); 3, BamHI-SstI

15   fragment containing the lactose operon of E.

coli without promoter, operator, translation

initiation site and the first 8 non-essential

codons of lacZ (Casadaban et al., 1983); 4,

Sau3AI fragment of Pseudomonas putida KT2440

20   (Bagdasarian et al., 1981); ori, origin of

replication; nic, relaxation site; Km',

kanamycin resistance gene; Sm', streptomycin

resistance gene; mob, locus essential for

mobilisation (Bagdasarian et al., 1981); B,

25   BamHI; C, ClaI; E, EcoRI; H, HindIII; P, PstI;

S, SstI; Sa, SalI; X, XhoI; Xb, XbaI.

Figure 31:   Retention time of different corrinoid

standards (1 mg/standard) on the separation

system described in Example 7. The column used

is a Nucleosil C-18 column (Macherey-Nagel).

5   Against each absorbance peak, a number is

shown corresponding to the corrinoid described

below. The retention time is shown as abscissa

and the absorbance at 371 nm appears as

ordinate.

10   1, cobyrinic acid; 2, cobyrinic acid a-amide;

3, cobyrinic acid g-amide; 4, cobyrinic acid

a,g-diamide; 5, cobyrinic acid c-amide; 6,

cobyrinic acid c,g-diamide; 7, cobyrinic acid

a,c-diamide; 8, cobyrinic acid triamide; 9,

15   cobyrinic acid tetraamide; 10, cobyrinic acid

pentaamide; 11, cobyric acid; 12, GDP-

cobinamide; 13, cobinamide phosphate; 14,

cobinamide; 15, cyanocobalamin 5'-phosphate;

16, cyanocobalamin.

20   Figure 32:   Nucleotide sequence of both strands of the

4748-bp SalI-SalI-SalI-SalI-SalI-BglI fragment

of Pseudomonas denitrificans. The strand

situated at the top is to be read from 5' to

3' in the left-to-right direction which

25   corresponds to the left-to-right orientation

of the fragment of the restriction map

presented in Figure 6.

Figure 33: Nucleotide sequence of both strands of the

3855-bp SstI-SstI-BamHI fragment of

Pseudomonas denitrificans. The strand situated

at the top is to be read from 5' to 3' in the

5   left-to-right direction which corresponds to

the left-to-right orientation of the fragment

of the restriction map presented in Figure 6.

Figure 34: Analysis of the probabilities of the coding

frames on the basis of codon preference using

10   the programme of Staden and MacLachlan (1982)

on the six reading frames of the 4748-bp SalI-

SalI-SalI-SalI-SalI-BglI fragment. For the

frames belonging to the same coding strand,

the most probable frame corresponds to that in

15   which a dotted line, not interrupted by stop

codons, is placed under the probability line

for this frame. 4a. Analysis of the sequence

corresponding to nucleotides 200 to 800. This

analysis enables open reading frame 14 to be

20   identified. It begins at the ATG at position

660 and ends at the TGA at position 379. 4b.

Analysis of the sequence corresponding to

nucleotides 800 to 1500. This analysis enables

open reading frame 15 to be identified. It

25   begins at the GTG at position 925 and ends at

the TAA at position 1440. 4c. Analysis of the

sequence corresponding to nucleotides 1450 to

2600. This analysis enables open reading frame 16 to be identified. It begins at the ATG at position 1512 and ends at the TGA at position 2510. 4d. Analysis of the sequence

5   corresponding to nucleotides 2500 to 4650. This analysis enables open reading frame 17 to be identified. It begins at the GTG at position 2616 and ends at the TGA at position 4511.

10   Figure 35:   Analysis of the probabilities of the coding frames on the basis of codon preference using the programme of Staden and MacLachlan (1982) on the six reading frames of the 3855-bp SstI-SstI-BamHI fragment. For the frames belonging

15   to the same coding strand, the most probable frame corresponds to that in which a dotted line, not interrupted by stop codons, is placed under the probability line for this frame. 5a. Analysis of the sequence

20   corresponding to nucleotides 1 to 905. This analysis enables open reading frame 18 to be identified. It begins at the ATG at position 809 and ends at the TGA at position 108. 5b. Analysis of the sequence corresponding to

25   nucleotides 955 to 2105. This analysis enables open reading frame 19 to be identified. It begins at the ATG at position 1971 and ends at

the TGA at position 1063. 5c. Analysis of the
sequence corresponding to nucleotides 2000 to
3300. This analysis enables open reading frame
20 to be identified. It begins at the ATG at

5   position 2099 and ends at the TAG at position
3115. 5d. Analysis of the sequence
corresponding to nucleotides 3250 to 3855.
This analysis enables open reading frame 21 to
be identified. It begins at the ATG at

10   position 3344 and ends at the TGA at position
3757.

Figure 36:   Construction of plasmids pXL233, pXL843 and
pXL1558 from pXL154.

The plasmids are constructed in the following

15   manner. The 3.5-kb EcoRI fragment containing
the truncated cobS gene and the sequence
upstream is excised from pXL154, then purified
and cloned to the EcoRI site of pKT230. The
plasmid thereby constructed is designated   .

20   pXL233. The 3.5-kb EcoRI-XhoI-XhoI fragment
containing the cobT gene and the sequence
downstream is excised and purified from pXL154
by partial digestions. The 4.3-kb EcoRI-EcoRI-
EcoRI fragment containing the cobS gene and

25   the sequence upstream is excised and purified
from pXL154 and then ligated to the above 3.5-
kb fragment. The approximately 8-kb EcoRI-XhoI

fragment thereby attained is cloned into the EcoRI and SalI sites of pXL59 to generate plasmid pXL843. Plasmid pXL1558 is constructed in the following manner: the 12-kb HindIII-HindIII fragment is excised from pXL154 and purified, and the ends are then filled in with the large fragment of E. coli DNA polymerase I. This insert is cloned in pRK290 (Ditta et al., 1981) digested with EcoRI and then treated with the large fragment of E. coli DNA polymerase I in order to make the ends blunt. Restriction sites which are shown in brackets correspond to sites which have disappeared during cloning. 1, PstI-SstI fragment of RSF1010 (Degraff et al., 1978); 2, PstI-BamHI fragment of pACYC177 (Bagdasarian et al., 1981); B, BamHI; C, ClaI; E, EcoRI; H, HindIII; P, PstI; S, SstI; Sa, SalI; X, XhoI; Xb, XbaI; Tet tetracycline resistance gene; Km', kanamycin resistance gene; Sm', streptomycin resistance gene.

Figure 37: Study of the insertions of the transposon Tn5Sp into the 12-kb HindIII-HindIII insert of pXL154.

The insertions of the transposon are mapped on the 12-kb HindIII-HindIII insert cloned into

pXL1558. The chromosomal insertions into
strain SC510 Rif' are boxed, that which is not
is introduced into strain SBL27 Rif'. A plus
or minus sign is shown under each insertion to

5   indicate the Cob phenotype of the strain
having this insertion. Absence of
complementation (or complementation) of strain
G2035 by plasmids pXL1558::Tn5Sp is indicated
by minus (or plus) signs below each insertion.

10   The inserts of the plasmids described in
Figure 36 are shown. The plus (or minus) signs
over these plasmids, and aligned with the
transposon insertions, show diagrammatically
the complementation (or absence) of the

15   transposon-mutated strain by the plasmid. The
open reading frames deduced from the sequence
are also given in this figure (ORF14 to 17, as
well as the corresponding cob genes (cobS and
cobT)). E : EcoRI; H : HindIII; X : XhoI.

20   Figure 38:   Construction of plasmids pXL1286, pXL1303,
pXL1324, pXL1490B and pXL1557 from pXL519.
The position of the sequenced fragment appears
in the upper part of the figure above the
restriction map of the cluster; it is a 3.9-kb

25   SstI-SstI-SstI-BamHI fragment. The plasmids
are constructed in the following manner. The
2-kb BglII-EcoRI fragment containing the cobU

gene and the sequence downstream is excised
from pXL519, then purified and cloned at the
BamHI and EcoRI sites of pKT230 to generate
plasmid pXL1286. The 2.7-kb SstI-EcoRI
fragment containing the truncated cobV gene,
cobU gene and the sequence downstream is
excised on pXL519, then purified and cloned at
the SstI and EcoRI sites of pKT230 to generate
plasmid pxL1324. The 1.6-kb SstI-SstI fragment
containing the truncated cobV gene and the
sequence upstream is excised from pXL519, then
purified and cloned at the SstI site of pKT230
to generate plasmid pXL1303. The 3.85-kb SstI-
SstI-BamHI fragment is purified after total
digestion of pXL519 with BamHI and partial
digestion with SstI. This fragment is then
cloned at the BamHI and SstI sites of pKT230
to generate pXL1490B. Plasmid pXL1557 is
constructed in the following manner: the 9-kb
HindIII-BamHI fragment is excised from pXL519
and purified, and the ends are then filled in
with the large fragment of E. coli DNA
polymerase I. This insert is cloned into
pRK290 (Ditta et al., 1981) digested with
EcoRI and then treated with the large fragment
of E. coli DNA polymerase I to make the ends
blunt. The restriction sites which are shown

in brackets correspond to sites which have disappeared during cloning. 1, PstI-SstI fragment of RSF1010 (Degraff et al., 1978); 2, PstI-BamHI fragment of pACYC177 (Bagdasarian et al., 1981); B, BamHI;Bg, BglII; C, ClaI; E, EcoRI; H, HindIII; P, PstI; S, SstI; Sa, SalI; X, XhoI; Xb, XbaI; Tet', tetracycline resistance gene; Km' kanamycin resistance gene; Sm', streptomycin resistance gene.

Figure 39: Study of the insertions of the transposon Tn5Sp into the 9-kb HindIII-BamHI insert of pXL519. The insertions of the transposon are mapped on the 9-kb HindIII-BamHI insert cloned into pXL1557. The chromosomal insertions into strain SC510 Rif' are boxed, those which are not are introduced into strain SBL27 Rif'. A plus or minus sign is shown under each insertion to indicate the Cob phenotype of the strain having this insertion. Absence of complementation (or complementation) of strain G2040 by plasmids pXL1557::Tn5Sp is indicated by minus (or plus) signs below each insertion. The inserts of the plasmids described in Figure 6 are shown. The plus (or minus) signs over these plasmids and aligned with the transposon insertions, show diagrammatically the complementation (or absence) of the

transposon-mutated strain by the plasmid. The
open reading frames deduced from the sequence
are also given in this figure (ORF18 to 21),
as well as the corresponding cob genes (cobU

5   and cobV).

Figure 40:   Coding sequences of each of the genes of the
4.8-kb fragment, cobX, cobS and cobT,
respectively, are indicated. The sequence of
the COBX, COBS and COBT proteins encoded by

10   these sequences appears under the respective
coding sequences cobX, cobS and cobT. The
legend is identical to that for Figure 15.

Figure 41:   Coding sequences of each of the genes of the
3.9-kb fragment, cobU and cobV, respectively,

15   are indicated. The sequence of the COBU and
COBV proteins encoded by these sequences
appears under the respective coding sequences
cobU and cobV. The legend is identical to that
of Figure 15.

20   Figure 42:   A. Total proteins of the strains E. coli BL21
pLysS pET3b and E. coli BL21 pLysS pXL1937
analysed in 10% SDS-PAGE. Lane 1, BL21
pLyspET3b; lane 2, E. coli BL21 pLysS pXL1937.
B. Total proteins of the strains E. coli BL21,

25   E. coli BL21 pXL1874 and E. coli BL21 pXL1875
analysed in 10% SDS-PAGE. Lane 1, E. coli

BL21; lane 2, <u>E. coli</u> BL21 pXL1874; lane 3, <u>E.</u> <u>coli</u> BL21 pXL1875.

The molecular masses of the markers are indicated. The band corresponding to the overexpressed protein is indicated by an arrow.

Figure 43:   Nucleotide sequence of both strands of the 13144-bp <u>Sst</u>I-<u>Sst</u>I-<u>Sst</u>I-<u>Sst</u>I-<u>Bgl</u>II-<u>Bgl</u>II fragment of <u>Pseudomonas</u> <u>denitrificans</u>. The strand situated at the top is to be read from 5' to 3' in left-to-right direction which corresponds to the left-to-right orientation of the fragment of the restriction map presented in Figure 46.

Figure 44:   Restriction map of the 13144-bp <u>Sst</u>I-<u>Sst</u>I-<u>Sst</u>I-<u>Sst</u>I-<u>Bgl</u>II-<u>Sst</u>I-<u>Bgl</u>II fragment of <u>Pseudomonas</u> <u>denitrificans</u>. The position or positions of restriction sites occurring are indicated in increasing order of the cut number on the fragment sequenced; the positions correspond to the sequence presented in Figure 43.

Figure 45:   Analysis of the probabilities of the coding frames on the basis of codon preference using the programme of Staden and MacLachlan (1982) on the six reading frames of the 13144-bp <u>Sst</u>I-<u>Sst</u>I-<u>Sst</u>I-<u>Sst</u>I-<u>Bgl</u>II-<u>Sst</u>I-<u>Bgl</u>II fragment

of <u>Pseudomonas</u> <u>denitrificans</u>. For the frames
belonging to the same coding strand, the most
probable frame corresponds to that in which a
dotted line, not interrupted by stop codons,
is placed under the probability line for this
frame.

1. Sequence corresponding to nucleotides 1 to
2266. This analysis enables open reading frame
22 to be identified. It begins at the ATG at
position 429 and ends at the TAG at position
1884.

2. Sequence corresponding to nucleotides 2266
to 4000. This analysis enables open reading
frame 23 to be identified. It begins at the
ATG at position 3364 and ends at the TGA at
position 3886.

3. Sequence corresponding to nucleotides 3800
to 5000. This analysis enables open reading
frame 24 to be identified. It begins at the
ATG at position 3892 and ends at the TAG at
position 4954.

4. Sequence corresponding to nucleotides 5000
to 9000. This analysis enables open reading
frame 25 to be identified. It begins at the
ATG at position 5060 and ends at the TAG at
position 8885.

5. Sequence corresponding to nucleotides 9000 to 9700. This analysis enables open reading frame 26 to be identified. It begins at the ATG at position 9034 and ends at the TGA position 9676.

6. Sequence corresponding to nucleotides 9600 to 13144. This analysis enables open reading frames 27, 28, 29 and 30 to be identified. They begin, respectively, at the ATGs at positions 9678, 10895, 11656 and 13059, and end at the stop codons at positions 10101, 10304, 12181 and 12366. Open reading frames 28 and 30 occur on the strand complementary to the coding strand corresponding to all the other open reading frames.

Figure 46: 13.4-kb EcoRI-BglII-EcoRI-BglII fragment, positions of the insertions of transposons Tn5Sp into the 9.1-kb EcoRI fragment, positions of the insertions of transposons Tn5 into the insert of plasmid pXL189 as well as the inserts of the various plasmids used during the experiments on complementation of strains SC510 Rif'::Tn5Sp. The complementations of the mutants SC510 Rif' ::Tn5Sp by the plasmids are indicated (+) - between 5% and 100% of the level of the parent strain SC510 Rif'-, (±) - partial

complementation, between 0.5 and 5% of the
level of SC510 Rif'-, or (-) - absence of
complementation, i.e. less than one thousand
times less than SC510 Rif'-, positioned

immediately above the lines showing
diagrammatically the insert of the plasmids
and aligned with the insertion sites of the
corresponding mutants. Below the mapping of
the insertions of transposons Tn5 into the

insert of plasmid pXL189, the complementation
(+) or absence of complementation (-) of these
mutant plasmids for the Agrobacterium
tumefaciens mutants G632 and G633 is shown. On
the right-hand part of the figure, there is a

table showing the complementation of the
mutants G622, G623 and G630 (Cameron et al.,
1989) by different plasmids; (+) - total
complementation, 100% of the level of the
parent strain C58C9 Rif'-, (±) - partial

complementation, between 10 and 50% of the
level of C58C9 Rif'-, or (-) - absence of
complementation.

The different plasmids whose insert is shown
are constructed as follows (the fragments are
excised either from pXL156 or from pXL157):

pXL618 corresponds to the 2.5-kb EcoRI-BamHI fragment cloned at the same sites of pKT230 (Bagdasarian et al., 1981);

pXL593 corresponds to the 3.1-kb BamHI fragment cloned at the BamHI site of pKT230 (Bagdasarian et al., 1981);

pXL623 corresponds to the 1.9-kb BamHI-XhoI fragment cloned at the BamHI-SalI sites of pXL59 (Cameron et al., 1989);

pXL1909 corresponds to the 8.4-kb BamHI-BamHI-BamHI fragment cloned at the BamHI of pKT230 (Bagdasarian et al., 1981);

pXL221 corresponds to the 1.6-kb EcoRI-ClaI fragment cloned at the same sites of pXL59 (the ClaI site into which this fragment is cloned is the ClaI site of the multisite of pXL59) (Cameron et al., 1989);

pXL1908 and 1938 correspond to the same insert, 6.5-kb XhoI-BamHI-BamHI fragment, to which XbaI linkers have been added; this insert is cloned in both orientations at the XbaI site of pXL435 (Cameron et al., 1989); an arrow positioned on the figure indicates the position of the kanamycin resistance gene with respect to the ends of the insert of the two plasmids;

pXL208 corresponds to the 5.2-kb BamHI

fragment cloned at the BamHI site of pKT230

(Bagdasarian et al., 1981);

pXL297 corresponds to the 9.1-kb EcoRI

5   fragment cloned at the EcoRI site of pKT230

(Bagdasarian et al., 1981).

The open reading frames (ORF) defined by the

sequencing of the fragment (ORF 22 to 30) are

shown, as well as the corresponding cob genes;

10   an arrow indicates the polarity of the

transcription.

E, EcoRI; B, BamHI; Bg, BglII; Cl, ClaI; Sau,

Sau3AI; X, XhoI;

Figure 47:   Coding sequences of each of the genes of the

15   13.4-kb fragment, cobQ, cobP and cobW, cobN

and cobO, respectively, are indicated. The

sequences of the COBQ, COBP, COBW, COBN and

COBO proteins encoded by these sequences

appear under their respective coding sequence

20   cobQ, cobP, cobW, cobN and cobO. The legend

is identical to that for Figure 15.

Figure 48:   A - NH_2-terminal sequence of SUMT of M.

ivanovii and sequence of the oligonucleotides

923, 946, 947; -, means that, at this

25   position, the residue could not be determined;

for the antisense oligonucleotide, the amino

acids indicated below the sequence correspond

to the anticodons shown. B - Presentation of
the enzymatic amplification of a fragment
internal to the structural gene of SUMT of M.
ivanovii with the oligonnucleotides 946 and
947.

Figure 49:   Construction of the recombinant replicative
form pG10. The 615-bp fragment obtained by
amplification is digested with HindIII and
EcoRI and then purified as described. This
fragment is then ligated with the replicative
form of phage M13mp19 digested with the same
enzymes. The recombinant clone is found as
described in the text.

Figure 50:   Autoradiograph of a genomic DNA blot of M.
ivanovii digested with various enzymes,
separated by agarose gel electrophoresis and
then transferred onto a nylon membrane as
described previously. The membrane is
hybridised with the pG10 probe as described
previously. 1, HindIII-BglII; 2, KpnI-BglII;
3, EcoRI-BglII; 4, BglII-PstI. The sizes of
the different fragments which hybridise with
the probe are shown in kb.

Figure 51:   Nucleotide sequence of both strands of the
955-pb fragment of M. ivanovii. The strand
situated at the top is to be read from 5' to
3' in the left-to-right direction.

Figure 52: Coding sequence of the corA gene of M. ivanovii obtained from the 955-bp sequence. The primary sequence of the CORA protein is also shown. The amino acids are shown above

5 their codon and the stop codon is designated by a star. The main physical properties of the CORA protein of M. ivanovii, namely the amino acid composition, in number and in percentage, the molecular weight, the index of polarity,

10 the isoelectric point and the optical density at 280 nm of a solution containing 1 mg/l of purified protein. The hydrophobicity profile of the CORA protein of M. ivanovii; this profile was obtained on the basis of the

15 programme of Hopp and Woods (1981). Positive values correspond to regions of the protein which are hydrophilic. The position of the amino acids is indicated as abscissa, and the value of the index of hydrophilicity as

20 ordinate; when this value is positive, this indicates that the protein is hydrophilic in this region.

Figure 53: Comparison of the primary sequences of the proteins COBA of P. denitrificans and CORA of

25 M. ivanovii. The proteins have been aligned by means of the programme of Kanehisa (1984). =, identical amino acids; -, homologous amino

acids on the basis of the criteria defined
above (see Figures 22 and 23).

Figure 54:   Construction of plasmids pXL1832 and pXL1841.
The legends described, placed on the figure,
enable the constructions to be followed.

**General techniques of cloning, molecular biology and
biochemistry.**

The classical methods of molecular biology, such as
centrifugation of plasmid DNA in a caesium
chloride/ethidium bromide gradient, digestions
with restriction enzymes, gel electrophoresis,
electroelution of DNA fragments from agarose
gels, transformation in E. coli, and the like,
are described in the literature (Maniatis et
al., 1982, Ausubel et al., 1987).

Restriction enzymes were supplied by New-England
Biolabs (Biolabs), Bethesda Research
Laboratories (BRL) or Amersham Ltd (Amersham).
Linker oligonucleotides were supplied by
Biolabs.

For the ligations, the DNA fragments are separated
according to their size on 0.7 % agarose or 8 %
acrylamide gels, purified by electroelution,
extracted with phenol, precipitated with ethanol
and then incubated in 50 mM Tris-HCl buffer pH
7.4, 10 mM $MgCl_2$, 10 mM DTT, 2 mM ATP, in the
presence of phage T4 DNA ligase (Biolabs).

If necessary, DNA fragments having protuberant 5' ends are dephosphorylated by a treatment with calf intestinal alkaline phosphatase (CIP, Pharmacia) at 37°C for 30 min in the following buffer:

5   100 mM glycine, 1 mM $MgCl_2$, 1 mM $ZnCl_2$, pH 10.5. The same technique is used for dephosphorylation of protuberant or blunt 3' ends, but the treatment is for 15 min at 37°C and then 15 min at 56°C. The enzyme is inactivated by heating

10   the reaction mixture to 68°C for 15 min in the presence of 1 % SDS and 100 mM NaCl, followed by a phenol/chloroform extraction and an ethanol precipitation.

Filling-in of protuberant 5' ends is performed with the

15   Klenow fragment of _E. coli_ DNA polymerase I (Biolabs). The reaction is performed at room temperature for 30 min in 50 mM Tris-HCl buffer pH 7.2, 0.4 mM dNTPs, 10 mM $MgSO_4$, 0.1 mM DTT, 50 mg/ml BSA. Filling-in of protuberant 3' ends is performed

20   in the presence of phage T4 DNA polymerase (Biolabs) according to the manufacturer's recommendations. Digestion of the protuberant ends is performed by limited treatment with S1 nuclease (BRL) according to the manufacturer's recommendations. Linker

25   oligonucleotides are added onto ends of DNA fragments as already described (Maniatis, 1982).

In vitro mutageneis with oligodeoxynucleotides is performed according to the method developed by Taylor et al., 1985, using the kit distributed by Amersham.

5 The ligated DNAs are used for transforming the strain rendered competent: E. coli MC 1060 [D(lacIOPZYA)X74, galU, galK, strA', hsdR] for plasmids or E. coli TG1[D(lac proA,B), supE, thi, hsdD5/ F' traD36, proA*, B*, lacIᵠ,

10 lacZDM15] for replicative forms of phages derived from bacteriophage M13.

Plasmid DNAs are purified according to the technique of Birnboim and Doly, 1979. Minipreparations of plasmid DNA are made according to the protocol

15 of Klein et al., 1980. Preparations of chromosomal DNA of Gram-negative bacteria are produced as already described (Cameron et al., 1989).

Radioactive probes are prepared by nick translation

20 according to the method already detailed (Rigby et al., 1977). Hybridisations between DNA sequences as well as the immobilisation of nucleic acids on nitrocellulose membranes are performed as already described (Cameron et al.,

25 1989). In clonings for which there is a small probability of finding the desired recombinant clone, the latter are found after hybridisation

on filters as already described (Maniatis et al., 1982).

The nucleotide sequence of DNA fragments is determined by the chain-termination method (Sanger et al., 1977). In the reaction mixture, dGTP is replaced by 7-deaza-dGTP, in order to avoid compression of bands during acrylamide gel electrophoresis caused by the high percentage of GC in the DNA.

The culture media used for the bacteriological part have already appeared (Maniatis et al., 1982). Culturing in PS4 medium is carried out as already described (Cameron et al., 1989); Pseudomonas denitrificans strains SC510 Rif' and G2 Rif' are cultured in PS4 medium as follows: 250-ml Erlenmeyers containing PS4 medium (25 ml), with, if necessary, the selective antibiotic for the plasmid carried by each strain, are inoculated with a 1/100 dilution of saturated preculture in L medium (Miller 1972), with, if necessary, the selective antibiotic for the plasmid carried by each strain; these cultures are incubated for 6 days at 30°C and the musts are then analysed for their cobalamin content or alternatively the enzymatic activity of some enzymes of the pathway. Strains of Agrobacterium tumefaciens, Pseudomonas putida and Rhizobium meliloti are cultured at 30°C;

except where otherwise stated, they are cultured

in L medium.

Bacterial conjugations are carried out as already

described (Cameron et al., 1989).

5   Extracts of total proteins are produced as already

described (Ausubel et al., 1987).

Analytical electrophoresis (SDS-PAGE) of proteins in

acrylamide gel under denaturing conditions is

performed as already described (Ausubel et al.,

10   1987). The PhastSystem apparatus (Pharmacia) using

Laemli's discontinuous-buffer system (Laemli, 1970)

is also used; different gels are used in accordance

with the molecular weights of the proteins to be

analysed as well as their purity:

15   PhastGel gradient 8-25

PhastGel Homogeneous 12.5

Staining is performed either with Coomassie blue with the

aid of PhastGel Blue R (Pharmacia), or with silver nitrate

using the PhastGel silver Kit (Pharmacia) in accordance

20   with the manufacturer's instructions.

NH_2-terminal sequences of the proteins are determined by

the Edman degradation technique, using an automated

sequencer (Applied Biosystems model 407A) coupled to

an HPLC apparatus for identification of the

25   phenylthiohydantoin derivatives.

EXAMPLE 1 - Isolation of DNA fragments of P. denitrificans
containing Cob genes

This example describes the isolation of DNA
fragments of Pseudomonas denitrificans carrying Cob genes.

5   These fragments were demonstrated by complementation
experiments on Cob mutants of A. tumefaciens and P. putida
(Cameron et al., 1989).

These Cob mutants were obtained by mutagenesis
with N-methyl-N'-nitro-N-nitrosoguanidine according to the

10   technique of Miller (Miller et al., 1972), or by insertions
of transposon Tn5. In this manner, strains incapable of
synthesising cobalamins were demonstrated, and especially
the Cob mutant G572 of P. putida and the Cob mutants G159,
G161, G164, G169, G171, G258, G609, G610, G611, G612, G613,

15   G614, G615, G616, G620, G622, G623, G630, G632, G633, G634,
G638, G642, G643, G2034, G2035, G2037, G2038, G2039, G2040,
G2041, G2042 and G2043 of A. tumefaciens.

At the same time, a library of genomic DNA of P.
denitrificans is produced in a mobilisable broad host range

20   vector pXL59, by digestion of 5 μg of DNA in the presence
of restriction enzymes (Cameron et al., 1989).

By complementation, several plasmids could be
isolated, enabling the Cob mutants of P. putida and of A.
tumefaciens to be complemented. Among these, plasmids

25   pXL151, pXL154, pXL156, pXL157 and pXL519 will be noted
more especially.

These plasmids were isolated and DNA fragments
could be excised, purified and analysed by restriction.
These fragments are presented in Figures 6 and 44: a 5.4-
kb ClaI-HindIII-HindIII-HindIII fragment, an 8.7-kb EcoRI-

5   EcoRI fragment, a 4.8-kb SalI-SalI-SalI-SalI-SalI-BglI
fragment, a 3.9-kb SstI-SstI-BamHI fragment and a 13.4-kb
EcoRI-BglII-EcoRI-BglII fragment.

EXAMPLE 2 - Sequencing of the DNA fragments isolated

        This example illustrates the sequencing of DNA

10   fragments carrying cob genes of Pseudomonas denitrificans
SC510.

        2.1. Sequencing of a 5.4-kb ClaI-HindIII-
HindIII-HindIII fragment.

        This fragment is contained in plasmid pXL157

15   described in Example 1. After excision, the subfragments of
the 5.4-kb fragment were cloned into phages M13mp18 or
M13mp19 (Norrander et al., 1983) or M13tg130 or M13tg131
(Kieny et al., 1983) in both orientations. Deletions were
then produced in vitro by the method of Henikoff (1987).

20   These deletions were then sequenced with the "universal
primer" as a synthetic primer of chain-termination
reactions. The overlap between these different deletions
enabled the total sequence, over both strands, of the
5.4-kb fragment to be established (Figure 7). This fragment

25   comprises 5378 bp. In the sequence described in Figure 7,
there are seen, before the ClaI site, three restriction
sites (PstI, SalI and XbaI) which have appeared during the

cloning of the fragment in question with a view to sequencing in cloning multisites. When subsequent reference is made, in the present invention, to the sequence of this ClaI-HindIII-HindIII-HindIII fragment, this will be to the

5   sequence presented in Figure 7 in which the first 22 bases do not correspond to the DNA of Pseudomonas denitrificans (thus, all the positions of restriction site or of beginning of open reading frame refer to the sequence presented in Figure 7).

10   2.2. Nucleotide sequence of an 8.7-kb EcoRI-EcoRI fragment.

      This fragment is carried by pXL151 described in Example 1. The EcoRI site as well as the adjacent 70 bp located to the right of this fragment originate from pXL59,

15   which is the vector used for constructing pXL151 by cloning an Sau3AI fragment of Pseudomonas denitrificans SC510. After excision, subfragments of the 8.7-kb fragment were cloned into phages M13mp18 or M13mp19 (Norrander et al., 1983) or M13tg130 or M13tg131 in both orientations (Kieny

20   et al., 1983). Deletions were then produced in vitro by the method of Henikoff (1987). These deletions were then sequenced with the "universal primer" as a synthetic primer of chain-termination reactions. The overlap between these different deletions enabled the total sequence, over both

25   strands, of the 8.7-kb fragment to be established (Figure 8). This fragment comprises 8753 bp.

2.3. Sequencing of a 4.8-kb SalI-SalI-SalI-SalI-SalI-BglI fragment.

This fragment is contained in plasmid pXL154 described in Example 1. The protocol is identical to that used in Example 2.2. The total sequence on both strands of the 4.8-kb fragment is presented in Figure 32. This fragment contains 4749 bp.

2.4. Nucleotide sequence of a 3.9-kb SstI-SstI-BamHI fragment.

This fragment is included in plasmid pXL519 described in Example 1. The protocol is identical to that used in Example 2.2. The total sequence on both strands of the 3.9-kb fragment is presented in Figure 33. This fragment contains 3855 bp.

2.5. Nucleotide sequence of a 13.4-kb EcoRI-BglII-EcoRI-BglII fragment.

This fragment is contained in plasmids pXL156 and pXL157 described in Example 1. The protocol used is identical to that of Example 2.2. The sequence on both strands of the 13.15-kb fragment is presented in Figure 43. It corresponds to the total sequence of the 13.4-kb fragment except for 250 bp, corresponding to an EcoRI-SstI fragment, occurring at the left-hand end of the fragment.

From these nucleotide sequences, restriction maps were obtained for the enzymes which cut least frequently (Figures 6 and 44). The percentage of GC bases in Pseudomonas denitrificans SC150 DNA is relatively high

(65.5 %) and manifests itself in compression on the sequencing gels. To avoid these problems, two approaches are adopted:

   i) the use of 7-deaza-dGTP instead of dGTP in the sequencing reactions to decrease the secondary structures which form during electrophoresis in the sequencing gel, and

   ii) the sequencing of both strands.

EXAMPLE 3 - Analysis of these nucleotide sequences: determination of the open reading frames

   The nucleotide sequences of the 5.4-kb ClaI-HindIII-HindIII-HindIII (Figure 7), 8.7-kb EcoRI-EcoRI (Figure 8), 4.8-kb SalI-SalI-SalI-SalI-SalI-BglI (Figure 32), 3.9-kb SstI-SstI-BamHI (Figure 33) and 13.4-kb EcoRI-BglII-EcoRI-BglII (Figure 43) fragments enable open reading frames to be defined. Since the DNA in question contains a high percentage of GC, the open reading frames are numerous in view of the low frequency of translation stop codons. A study of the probability of the coding frames on the basis of codon preference using the method of Staden and MacLachlan (1982) is carried out. It characterises the open reading frames which have the greatest probability of being coding relative to the other frames of the same DNA strand, this probability being dependent on the codon preference of genes already sequenced originating from bacteria of the genus Pseudomonas. In this manner:

3.1. Five open reading frames are characterised

for the 5.4-kb ClaI-HindIII-HindIII-HindIII fragment. They

are designated frames 1 to 5, and their positions in the

sequence of the 5.4-kb fragment are as follows (in the

5   5'→3' sequence from the ClaI site to the HindIII sites):

Table: Probable open reading frames of the 5.4-kb

ClaI-HindIII-HindIII-HindIII fragment. The positions in the

sequence correspond to the positions in the sequence

described in Figure 7; the coding strand is the 5'→3'

10   strand corresponding to the upper strand in this figure.

Frame number	Translation initiation codon	Stop codon	Molecular weight in kD of the encoded protein
1	549	1011	15.5
2	1141	1980	29.2
3	1980	3282	5.7
4	3281	4280	35.0
5	4284	5253	34.1

The representations of the probabilities

that these open reading frames are coding frames, with

those observed on the other frames (5 in total) in

30   parallel, are given in Figure 9. These five frames are

encoded by the same strand. Four of them (open reading

frames 1 to 4) display the characteristics of coding frames

in translational coupling (Normak et al., 1983), namely,

the translation initiation codon of frame x+1 overlaps the

translation termination codon of frame x, or else these
codons are very close.

       3.2. Eight frames are characterised for the
8.7-kb EcoRI-EcoRI fragment. They are designated frames 6
to 13 and their positions in the sequence of the 8.7-kb
fragment are given in the table below.

Table: Probable open reading frames of the 8.7-kb EcoRI
fragment. The positions in the sequence correspond to the
positions in the sequence described in Figure 8; in this
figure, the coding strand is the upper strand.

Translation frame number	Initiation codon	Stop codon	Molecular weight in kD of the encoded protein
6	736	1519	28.9
7	1620	2997	46.7
8	3002	3632	22.0
9	3631	4366	25.8
10	4365	5127	27.1
11	5126	5867	26.8
12	5862	7101	42.9
13	7172	7931	26.8

       The representations of the probabilities of
these open reading frames, with those observed on the other
frames (6 frames in total) in parallel, are given in Figure
10. With the exception of frame 11, these eight frames are
encoded by the same strand. Four of them (from 7 to 10)
display the characteristics of coding frames in

translational coupling (Normark et al., 1983), namely, the translation initiation codon of frame x+1 overlaps the translation termination codon of frame x, or else these codons are very close.

3.3. Four open reading frames are characterised for the 4.8-kb SalI-SalI-SalI-SalI-SalI-BqlI fragment. They are designated phases 14 to 17 and their positions in the sequence of the 4.8-kb fragment are as follows (in the 5' - 3' sequence from the SalI sites to the BqlI site):

Table: Probable open reading frames of the 4.8-kb SalI-SalI-SalI-SalI-SalI-BqlI fragment. The positions in the sequence corresponds to the positions described in Figure 32, where the upper strand is given in its 5' - 3' orientation. Frames 15, 16 and 17 are encoded by the upper strand, in contrast to frame 14.

Frame number	Translation initiation codon	Stop codon	Molecular weight in kD of the encoded protein
14	660	379	10286
15	925	1440	18941
16	1512	2510	36983
17	2616	4511	70335

The representations of the probabilities that these open reading frames are coding, with those observed on the other frames (4 in total) in parallel, are given in Figure 34. Frames 15, 16 and 17 are encoded by the same strand, frame 14 by the complementary strand.

3.4. Four frames are characterised for the 3.9-kb SstI-SstI-BamHI fragment. They are designated 18 to 21 and their positions in the sequence of the 3.9-kb fragment are given in the table below.

5   Table: Probable open reading frames of the 3.9-kb SstI-SstI-BamHI fragment. The positions in the sequence correspond to the positions described in Figure 33, where the polarity of the upper strand is 5' - 3'. Frames 18 and 19 are encoded by the lower strand, in contrast to frames

10   20 and 21.

Frame number	Translation initiation codon	Stop codon	Molecular weight in kD of the encoded protein
18	809	108	25148
19	1971	1063	30662
20	2099	3115	34682
21	3344	3757	14802

25   The representations of the probabilities that these open reading frames are coding, with those observed on the other frames (4 in total) in parallel, are given in Figure 35. Frames 19 and 20 are transcribed in a differing manner.

30   3.5. Nine open reading frames are characterised for the 13.1-kb EcoRI-BglII-EcoRI-BglII fragment. They are designated frames 22 to 30 and their positions in the sequence of the 13.1-kb fragment are as follows (in the 5' - 3' sequence from the EcoRI site to the BglII site):

Table: Probable open reading frames of the 13.1-kb
EcoRI-BglII-EcoRI-BglII fragment. The positions in the
sequence correspond to the positions described in Figure
43, where the upper strand is given in its 5' - 3'
orientation. The frames 22, 23, 24, 25, 26, 27 and 29 are
encoded by the upper strand, in contrast to the frames 28
and 30.

Translation frame number	Initiation codon	Stop codon	Molecular weight in kD of the encoded protein
22	429	1884	51982
23	3364	3886	19442
24	3892	4954	38121
25	5060	8885	138055
26	9034	9676	24027
27	9678	10101	14990
28	10835	10306	21057
29	11656	12181	19183
30	13059	12368	24321

The representations of the probabilities that
open reading frames 22, 23, 24, 25 and 26 are coding, with
those observed on the other frames (5 in total) in
parallel, are given in Figure 45. These 5 frames are
encoded by the same strand.

EXAMPLE 4 - Genetic studies on the DNA fragments carrying cob genes

This example shows the relationship which exists between the different open reading frames identified above and the genes involved in the biosynthesis of cobalamins and/or cobamides carried by these same fragments. These genes are identified by a genetic study as described below.

4.1 - Genetic study of the 5.4-kb fragment

Plasmid pXL723 is plasmid pRK290 (Ditta et al., 1980) containing the 2264-bp EcoRI-HindIII fragment corresponding to the right-hand portion of the fragment studied, cloned at the EcoRI site of pRK290 (Figure 11). The construction of the other plasmids used in this study (pXL302, pXL1397, pXL545, pXL545Ω, pXL556 and pXL1500) is described in the legend to Figures 11 and 12.

Insertions were obtained in plasmid pXL723 using the technique of de Bruijn and Lupski, 1984. Insertions of transposon Tn5 into plasmid pXL723 were selected and then mapped in the 5.4-kb fragment (Figure 12). pXL723 complements the Cob mutant G572 of Pseudomonas putida and the Cob mutant G634 of Agrobacterium tumefaciens. These insertions are classified in two groups of inactivating insertions: either those which no longer permit complementation of the Cob mutant G572, or those which abolish the complementation of the Cob mutant G634 (Figure 12). Insertions which inactivate the complementation of mutant G572 are mapped in open reading frame 4 (these are

insertions 15, 27, 68, 81 and 97); open reading frame 4

hence corresponds to a cob gene. The latter is designated

cobC. Insertions which inactivate the complementation of

mutant G634 are mapped in frame 5 (these are insertions 66

5   and 107, Figure 12); open reading frame 5 hence corresponds

to a cob gene. The latter is designated cobD. Moreover,

insertions with a transposon Tn5Sp' were produced.

Transposon Tn5Sp' was constructed in the laboratory by

cloning a BamHI cassette, containing the spectinomycin

10   resistance gene originating from plasmid pHP45Ω (Prentki

and Krisch, 1984), at the BamHI site of transposon Tn5

(Jorgensen et al., 1979). These insertions were made in the

chromosome of Pseudonomas denitrificans strain SBL27 Rif'.

Strain SBL27 is a strain of Pseudomonas denitrificans from

15   which SC510 is derived by several mutageneses. SBL27

produces 10-fold less cobalamins than SC510 on PS4 medium.

Of 10,000 clones of strain SBL27 Rif' each carrying an

insertion of transposon, more than 30 had lost the capacity

to synthesise cobalamins. Some of these clones possessed an

20   insertion in the fragment studied in this example. These

insertions were mapped by restriction analysis according to

Southern's method (Southern, 1975). The sites of insertions

of the transposon in these different mutants are given in

Figure 12. One of these insertions, number 2639, occurs in

25   the cobC gene; this insertion is complemented by plasmid

pXL302, which carries a fragment containing the cobC gene

(Figure 12). Two insertions, designated 2636 and 2638, are

in open reading frame 3. These mutants are blocked in the biosynthesis of cobalamins, and they are complemented by plasmid pXL1397 which contains only open reading frame 3, but non-complemented by plasmid pXL302 which contains the

5   cobC and cobD genes (Figure 12). Both of these insertions are hence in another gene. With open reading frame 3, we associate the cobB gene. An insertion 2933 is placed in open reading frame 2; it is complemented by plasmid pXL1500 which contains open reading frame 2; this insertion is non-

10   complemented by plasmid pXL1397, which contains the cobB gene and which complements the two insertions in cobB. In this instance, the insertion is hence in another gene; with open reading frame 2, we associate a gene designated cobA.

A kanamycin resistance cassette originating from

15   plasmid pUC4K (Barany et al., 1985) was introduced at the NotI site of the ClaI (position 0 in the sequence)-RsaI (position 1686 in the sequence) fragment cloned into a plasmid pUC8 (Viera and Messing, 1982); the NotI site in question is located at position 771 in frame 1 (see the

20   sequence in Figure 7); two insertions were adopted, each corresponding to a different orientation of the resistance cassette. These fragments, each carrying an insertion of the resistance cassette, were cloned into plasmid pRK404 (Ditta and al.) to give plasmids pXL1630 and 1631. These

25   plasmids were introduced by conjugative transfer into Pseudomonas denitrificans strain SC510 Rif', and then, by a series of cultures/dilutions in the absence of the

selective antibiotic for the plasmid (tetracycline), double recombinants which had exchanged the plasmid fragment with the chromosomal fragment and had lost the plasmid were found. Two strains were thereby characterised:

5   i) one is designated SC510:1631 Rif'; in this strain, the kanamycin resistance cassette is inserted in the chromosome at the NotI site (occurring in frame 1); the polarities of the transcriptions of the kanamycin resistance gene and that of open frame 1 are opposite,

10   ii) the other insertion is designated SC510:1630 Rif'; the resistance cassette is inserted at the same site, but the transcription of the resistance gene has the same polarity as that of the complete open reading frame 1.

These two strains both have a rate of synthesis of

15   cobalamins at least 100-fold lower than that of SC510.

Plasmid pXL545Ω corresponds to plasmid pXL545 into which the spectinomycin resistance cassette of plasmid pHP45Ω has been inserted at the BamHI site. This plasmid (Figure 12), which contains the 814-bp ClaI-HindIII

20   fragment (in which only open reading frame 1 is complete) complements only mutant SC510:1630 Rif'. This suffices to define a new gene, since this mutant is complemented by a plasmid which only contains the complete open reading frame 1. Open reading frame 1 corresponds to a gene of the

25   pathway of biosynthesis of cobalamins and/or cobamides. This gene is designated cobE. The absence of complementation of mutant SC510:1631 Rif' by plasmid

pXL545Ω is possibly due to the fact that the cobA, cobB, cobC, cobD and cobE genes, or a part of them, belong to the same operon, and that the insertion in cobE which preserves a transcription in the direction of transcription of the operon may be complemented only by trans expression of the cobE gene. In contrast, mutant SC510:1631 Rif', for its part, can be complemented only by a plasmid which permits trans expression of the cobA to cobE genes.

The 5.4-kb ClaI-HindIII-HindIII-HindIII fragment hence contains five cob genes designated cobA, cobB, cobC, cobD and cobE.

4.2 - Genetic studies of the 8.7-kb fragment

Plasmid pXL367 is pRK290 (Ditta et al., 1980) containing the 8.7-kb EcoRI fragment cloned at the EcoRI site (Figure 13).

Insertions of transposon Tn5 into plasmid pXL367 were selected using the technique already described (de Bruijn and Lupski, 1984). The insertions in the 8.7-kb fragment were mapped. In the same manner, insertions of transposon Tn3lacZ were obtained according to the method already described (Stachel et al., 1985) and mapped. 29 insertions of transposon Tn5 and 13 insertions of transposon Tn3lacZ were thus mapped. The precise position of these insertions in the 8.7-kb fragment is given in Figure 14. Plasmids each carrying a single insertion in the 8.7-kb fragment were introduced by conjugative transfers into the Cob mutants of Agrobacterium tumefaciens G164,

G609, G610, G611, G612, G613, G614, G615, G616, G620 and G638. These mutants are all complemented by pXL367. Insertions which no longer permit the complementation of the different mutants were sought. They correspond to an insertion in the gene responsible for complementation of the corresponding mutant. The results of the complementations of the different mutants for their character of production of cobalamins (Cob phenotype) are given in Figure 14. If the recombinant mutant produces less than threefold less cobalamins than are produced by the same mutant with plasmid pXL367, it is considered to be non-complemented. Of the mutants studied, G164, G609, G610, G611, G612, G613, G614, G615, G616, G620 and G638, eight different classes of inactivating insertions of transposons leading to a mutated phenotype are observed. These classes characterise insertions by the absence of complementation of one or more mutants by plasmids pXL367 carrying these same insertions. Each class hence corresponds to a mutated gene. It is observed that the insertions belonging to the same class are positioned beside one another. Eight classes of insertions are thus observed, which enable eight genes to be defined. Each class of insertions defines a minimum fragment which must be contained in the corresponding gene. Figure 14 demonstrates a perfect correlation between the regions bounded by each class, in respect of the restriction map, and the open reading frames described above (Example 3). It is found, in effect, that, for each

class of insertions, the transposons are always inserted in a portion of the 8.7-kb fragment which is contained in a single open reading frame. Each class of insertions is hence associated with one, and only one, open reading

5   frame. The open reading frames indicated above hence each code for a protein involved in the pathway of biosynthesis of cobalamins and/or cobamides. The open reading frames each correspond to genes involved in the biosynthesis of cobalamins and/or cobamides. These open reading frames are

10   referred to as cobF, cobG, cobH, cobI, cobJ, cobK, cobL and cobM for frames 6 to 13, respectively. The position of these genes relative to the restriction map is shown in Figure 14.

         4.3 - Genetic study of the 4.8-kb fragment

15   Plasmid pXL1558 is plasmid pRK290 (Ditta et al., 1980) containing the 12-kb HindIII-HindIII fragment of pXL154 (Cameron et al., 1989) cloned at the EcoRI site of pRK290 (Figure 36). The construction of the other plasmids used in this study (pXL233 and pXL843) is described in the

20   legend to Figure 36.

                Tn5Sp insertions were obtained in plasmid pXL1558. First, a strain containing a transposon Tn5Sp was constructed; this was done by transforming strain C2110 (Stachel et al., 1985) using plasmid pRK2013Tn5Sp (Blanche

25   et al., 1989); since it has a ColE1 origin of replication, plasmid pRK2013Tn5Sp does not replicate in strain C2110, which is polA-. The colonies obtained after transformation,

which are resistant to spectinomycin, hence have transposon

Tn5Sp in their chromosome; a colony is then reisolated,

after which the insertion of the transposon is then

transduced using phage Pl in strain MC1060 as described

5   previously (Miller, 1972). Strain MC1060 Tn5Sp is

transformed with plasmid pXL1558; plasmid pXL1558 is then

mobilised by conjugation using pRK2013 in C600 Rif'.

Conjugants resistant to tetracycline (for plasmid pXL1558)

and to spectinomycin (for the transposon) are then

10 selected. Such conjugants must contain plasmid pXL1558 in

which transposon Tn5Sp has been inserted. Insertions

carried in plasmid pXL1558, and more precisely in the 12-kb

fragment, are then mapped by restriction digestion; 23

insertions are thereby obtained and mapped on the 12-kb

15 fragment; the position of these different insertions in the

fragment is presented in Figure 37. These 23 insertions

were introduced into the chromosome of strain SC510 Rif'

after conjugative transfer of p-XL1558::Tn5Sp, followed by

introduction of plasmid pR751. Plasmid pR751 is a

20 trimethoprim-resistant plasmid of the same incompatibility

group as pXL1558 (incP, Thomas and Smith, 1987). By

culturing non-selectively for pXL1558 (absence of

tetracycline) but selectively for pR751 and the transposon

(presence of trimethoprim and of spectinomycin), the

25 exchange of the mutation carried by pXL1558::Tn5Sp with the

chromosome and also the segregation of pXL1558 are obtained

by the technique of marker exchange by double homologous

recombination, as already described (Schell et al., 1988).
The strains thereby selected carry the transposon in their
chromosome. The double homologous recombination is verified
by Southern's method (Southern, 1975). In this way, 23

5   SC510 Rif'::Tn5Sp strains in the 12-kb fragment were
identified.

Furthermore, another Tn5Sp insertion obtained by
random mutagenesis of transposon Tn5Sp in strain SBL27 Rif'
(Blanche et al., 1989) was mapped on the 12-kb fragment by

10   restriction analysis according to Southern's method
(Southern, 1975), see Figure 37; this strain is designated
SBL27 Rif'::Tn5Sp 1480.

The level of cobalamin synthesis is determined for
these 24 strains cultured in PS4 medium according to the

15   protocol already described (Cameron et al., 1989), and the Cob-
phenotype is assigned to strains producing at least 1000
(or 100) times less vitamin B_{12} than the parent strain SC510
Rif' (or SBL27 Rif'), Figure 37. It is thus observed that 6
of these chromosomal insertions lead to a Cob- phenotype

20   in P. denitrificans; they are the insertions 31.1, 41.3,
45, 55, 22.1 and 1480.

Three plasmids pxL233, pXL837 (Cameron et al.) and
pXL843 are introduced by conjugative transfers into three
strains possessing the Cob- phenotype, namely SC510

25   Rif'::Tn5Sp 31.1, SC510 Rif'::Tn5Sp 45 and SBL27 Rif'::Tn5Sp
1480. These three mutants each have a different
complementation profile for cobalamine synthesis. In

effect, SBL27 Rif'::Tn5Sp 1480 is complemented by pXL837 and pXL843 but not by pXL233; the mutant SC510 Rif'::Tn5Sp 45 is complemented only by pXL843; the mutant SC510 Rif'::Tn5Sp 31.1 is complemented by plasmid pXL843 and also by plasmid pXL233 (see Figure 37). The data presented hence enable it to be concluded, on the basis of the results of the complementations of the three mutants, that the three mutants are different and that, for each of them, transposon Tn5Sp has been inserted into a different cob gene.

Furthermore, plasmids pXL1558::Tn5Sp 41.3, pXL1558::Tn5Sp 45 and pXL1558::Tn5Sp 22.1 are introduced by conjugative transfers into strain G2035 (Cameron et al., 1989), and do not complement it. Plasmid pXL1558 complements this mutant, in contrast to plasmid pXL1558::Tn5Sp 31.1.

The phenotype and complementation data enable us to define 3 classes of insertions; each of these classes is represented by the following insertions : 31.1, class 1; 45, 41.3, 55 and 22.1, class 2; 1480, class 3.

For each class of insertions, the transposons are always inserted in a portion of the 4.8-kb fragment which is contained in a single open reading frame (ORF14, ORF16 and ORF17, as defined in Example 3). Each class of insertions is associated with a single open reading frame. The open reading frames indicated above hence code for a protein involved in the pathway of biosynthesis of

cobalamins and/or cobinamides. These open reading frames
are referred to as cobX, cobS and cobT for frames 14, 16
and 17. The position of these genes relative to the
restriction map is shown in Figure 37. Open reading frame

5   15 is not a gene involved in the biosynthesis of coenzyme
B_{12}.

4.4 - Genetic studies of the 3.9-kb fragment

Plasmid pXL1557 is plasmid pRK290 (Ditta et al.,
1980) containing the 9-kb HindIII-BamHI fragment of pXL519

10   cloned at the EcoRI site of pRK290 (Figure 38). The
construction of the other plasmids used in this study
(pXL1286, pXL1303, pXL1324) is described in the legend to
Figure 38. Moreover, the 2-kb BglII-XhoI fragment
(positions in the sequence presented in Figure 33: 251 and

15   2234) of plasmid pXL519 is cloned at the BamHI-SalI sites
of plasmid pXL435 (Cameron et al) to generate plasmid
pXL699.

Tn5Sp insertions were obtained in plasmid pXL1557
according to the technique described in Example 4.3.

20   Insertions of transposon Tn5Sp into plasmid pXL1557, then
designated pXL1557::Tn5Sp, were selected. Those which are
mapped in the 9-kb fragment (Figure 39) were introduced
into the chromosome of strain SC510 Rif' after conjugative
transfer of pXL1557::Tn5Sp and marker exchange by double

25   homologous recombination as described in 4.3.

The double homologous recombination is verified by Southern's method (Southern, 1975). In this way, 20 SC510 Rif':: Tn5Sp strains were identified.

Furthermore, two other Tn5Sp insertions obtained by random mutagenesis of transposon Tn5Sp in strain SBL27 Rif' (Blanche et al., 1989) were mapped on the 9-kb fragment by restriction analysis according to Southern's method (Southern, 1975), see the insertions 1003 and 1147 in Figure 39.

The level of cobalamin synthesis is determined for these 22 strains cultured in PS4 medium according to the protocol already described (Cameron et al., 1989), and the Cob- phenotype is assigned to strains producing 1000 (or 100) times less vitamin B_{12} than the parent strain SC510 Rif' (or SBL27 Rif'), Figure 39. Only the 4 insertions 1, 1003, 23 and 1147 result in a Cob- phenotype in P. denitrificans.

Four plasmids pXL699, pXL1286, pXL1303 and pXL1324 are introduced by conjugative transfers into the four strains possessing the cob- phenotype, namely SC510 Rif'::Tn5Sp 1, SBL27 Rif'::Tn5Sp 1003, SC510 Rif'::Tn5Sp 23 and SBL27 Rif'::Tn5Sp 1147. Plasmid pXL699 complements the first two mutants (SC510 Rif'::Tn5Sp 1, SBL27 Rif'::Tn5Sp 1003), but plasmid pXL1303 does not complement them, plasmid pXL1324 complements the other two mutants (SC510 Rif'::Tn5Sp 23 and SBL27 Rif'::Tn5Sp 1147) but plasmid pXL1286 does not complement them.

Furthermore, plasmid pXL1557::Tn5Sp 1, is
introduced by conjugative transfer into strain G2040, and
does not complement it, whereas plasmids pXL1557,
pXL1557::Tn5Sp 6A, pXL1557::Tn5Sp 54, pXL1557::Tn5Sp 48,
pXL1557::Tn5Sp 21, pXL1557::Tn5Sp 8, pXL1557::Tn5Sp 23,
also introduced by conjugative transfers, complement it
(see Figure 39).

The phenotype and complementation data enable 2
classes of insertions to be defined. For each class of
insertions, the transposons are always inserted in a
portion of the 3.9-kb fragment which is contained in a
single open reading frame (ORF19 and ORF20 as defined in
Example 3).

Each class of insertions is associated with a
single open reading frame. The open reading frames
indicated above code for a protein involved in the pathway
of biosynthesis of cobalamins and/or cobinamides. These
open reading frames are referred to as cobV and cobU for
frames 19 and 20. Frames 18 and 21 are not genes involved
in the pathway of biosynthesis of coenzyme B_{12}. The position
of these genes relative to the restriction map is shown in
Figure 39. The insertions 48, 21 and 8 are mapped between
the cobU and cobV genes.

4.5 - Genetic studies of the 13.4-kb fragment

4.5.1. Studies on the 4327-bp EcoRI-BglII fragment.

Plasmid pXL189 (Cameron et al., 1989), which
contains at least one cob gene, carries a 3.1-kb insert

which, except for 300 bp, corresponds to a 4.26-kb EcoRI-

ClaI fragment (see Figure 45). pXL189 was subjected to a

mutagenesis with transposon Tn5, as described previously

(De Bruijn and Lupski (1984)). 13 insertions were thereby

5   mapped in the insert of pXL189, as presented in Figure 46.

These 13 mutant plasmids, as well as pXL189, were

conjugated in two A. tumefaciens mutants, G632 and G633,

which are mutants complemented by pXL189 (Cameron et al.,

1989). Only the insertion 58 proved to be an inactivating

10   insertion. This result shows that the two mutants G632 and

G633 correspond to a mutation in the same gene, and that,

moreover, the only gene of P. denitrificans which could be

responsible for their complementation corresponds to open

reading frame 26 (see Figure 46), since insertion 58 is

15   mapped in this open reading frame; in addition, it is the

only insertion of the 13 which is mapped in this open

reading frame. A cob gene, designated cobO, is hence

associated with open reading frame 26.

To know whether the four open reading frames (open

20   reading frames 27 to 30) identified in this fragment

correspond to cob genes, a spectinomycin resistance

cassette from plasmid pHP45Ω (Prentki and Krisch, 1984) was

specifically inserted into each of these genes, and then

introduced into the chromosome of P. denitrificans SC510

25   Rif' by homologous recombination so as to obtain mutants of

insertions in each of these open reading frames. For this

purpose, the EcoRI-ClaI fragment (respective positions 8818

and 13082 in the sequence presented in Figure 43) was used.
This fragment, which carries the open reading frames 27 to
30, was purified from pXL157 (Cameron et al., 1989); an
EcoRI linker was added to the ClaI end after the latter had
5   been filled in with the Klenow fragment of E. coli DNA
polymerase. This fragment was then cloned into plasmid
pUC13 (Viera et al., 1982) at the EcoRI site. The plasmid
thus constructed was referred to as pXL332. Insertions of
the spectinomycin resistance cassette from plasmid pHP45Ω
10   (Prentki and Krisch, 1984) were carried out on pXL332.
These insertions were done separately at the SmaI (position
9868, open reading frame 27), BamHI (position 10664, open
reading frame 28), ClaI (position 11687, open reading frame
29) and NcoI (position 12474, open reading frame 30) sites
15   by total or partial digestions of pXL332 with the
corresponding enzymes, and then, if necessary, filling-in
of these ends with the Klenow fragment of E. coli DNA
polymerase, followed by ligation with the 2-kb SmaI
fragment of pHP45Ω (Prentki and Krisch, 1984) containing a
20   spectinomycin resistance gene; these insertions are
designated Ω2, Ω1, Ω3 and Ω4, respectively, as presented in
Figure 46. The EcoRI fragments carrying these different
insertions were then cloned into pRK404 (Ditta et al.,
1985) at one of the two EcoRI sites. The 4 plasmids
25   carrying these different insertions were then introduced by
conjugation in SC510 Rif', as described above. Plasmid
pR751 (Thomas and Smith, 1987) was then introduced into the

transconjugants. The exchange of mutations carried by the 4 different derivatives of pRK404 and the chromosome of SC510 Rif' could be selected as described (see Example 4.3). 4 strains were thereby obtained. These strains each carry an insertion of the resistance cassette in one of the four open reading frames 27 to 30. These insertions were verified by analysis of the genomic DNA by Southern blotting (Southern, 1975). The cobalamin production of these different strains was studied. They all showed a Cob+ phenotype on culturing in PS4 medium. This result indicates that these four open reading frames do not participate in the biosynthesis of coenzyme B_{12}. However, it is possible that one or more of these frames code for proteins which participate, e.g., in the conversion of coenzyme B_{12} to methylcobalamin for example, i.e. the synthesis of another cobalamin or even of another corrinoid.

4.5.2. Study of the 9.1-kb EcoRI-EcoRI fragment.

Various plasmids are used in this study; plasmid pXL1560 is plasmid pRK290 (Ditta et al., 1980) containing the 9.1-kb EcoRI-EcoRI fragment of pXL156 (Example 1) cloned at the EcoRI site of pRK290 (see Figure 46). The construction of the other plasmids used in this study (pXL618, pXL593, pXL623, pXL1909, pXL1938, pXL1908, pXL221, pXL208, pXL297) is described in the legend to Figure 45.

Tn5Sp insertions were obtained in plasmid pXL1560. Strain MC1060 Tn5Sp transformed with plasmid pXL1560 was used to obtain insertions of transposon Tn5Sp into the

pXL1560 fragment; 27 insertions were thereby obtained and mapped on the 9.1-kb fragment; the position of these different insertions in the fragment is presented in Figure 4. These 27 insertions were introduced into the chromosome of strain SC510 Rif' after conjugated transfer of pXL1560::Tn5Sp, followed by introduction of plasmid pR751. Plasmid pR751 is a trimethoprim-resistant plasmid of the same incompatibility group as pXL1560 (incP, Thomas and Smith, 1987). By culturing non-selectively for pXL1560 (absence of tetracycline) but selectively for pR751 and the transposon (presence of trimethoprim and of spectinomycin), the exchange of the mutation carried by pXL1560::Tn5Sp with the chromosome and also the segregation of pXL1560 are obtained; this technique of marker exchange by double homologous recombination is equivalent to that already described by Schell et al., 1988. The strains thus selected carry the transposon in their chromosome.

The double homologous recombination is verified by Southern's method (Southern, 1975). In this way, 27 SC510 Rif'::Tn5Sp strains each possessing a different insertion of transposon Tn5Sp in the 9.1-kb fragment were identified.

The level of cobalamin synthesis is determined for these 27 strains cultured in PS4 medium, and the Cob- phenotype is assigned to strains producing at least 1000 times less vitamin B_{12} than the parent strain SC510 Rif', Figure 46. It is thus observed that 18 out of the 27 of

these chromosomal insertions lead to a Cob- phenotype in P. denitrificans, as shown in Figure 46.

The insertions 19, 32, 24, 27, 37, 39, 26, 11 and 14 are mapped in open reading frame 22 (see Figure 46). All these insertions are complemented by plasmid pXL618, which contains only open reading frame 22. We deduce from this that open reading frame 22 corresponds to a cob gene, which we referred to as cobO. No insertion was obtained in open reading frame 23; however, plasmid pXL623, which contains only this open reading frame (see Figure 46), complements two cob mutants of Agrobacterium tumefaciens, G642 and G2043 (Cameron et al., 1989). Open reading frame 23 hence corresponds to a cob gene designated cobP. The insertions 23, 13, 12, 30, 22, 40, 35, 10 and 17 which are mapped in open reading frames 24 and 25 lead to a Cob- phenotype in SC510 Rif'. There hence appear to be two open reading frames whose product is involved in the biosynthesis of cobalamins. However, it cannot be ruled out that these insertions have polar effects on the genes positioned on the 3' side, such as cobO. It is hence appropriate to study the complementation of these mutants in order to determine whether the Cob- phenotype does not result from a polar effect.

The Cob mutants of Agrobacterium tumefaciens, G622, G623 and G630, complemented by pXL156, were studied. These mutants are not complemented by plasmid pXL189 (Cameron et al., 1989), which contains cobO as the only cob

gene. In contrast, they are complemented by plasmid
pXL1908, which contains cobO and open reading frame 25 in
addition to the open reading frames 27 to 30 (see Figure
45). The latter frames cannot be responsible for the
complementation of these mutants, since the proteins for
which they code do not participate in the coenzyme B_{12}
pathway. Hence, the observed complementations can only
result from open reading frame 25. In addition, the SC510
Rif' Tn5Sp mutants mapped in this same open reading frame
(these are the mutants 22, 40, 35, 10 and 17) are
complemented by plasmid pXL1908, see Figure 46, (carrying
cobO and frame 25), whereas at least two of them are not
complemented by pXL189, which contains only cobO as a cob
gene. These results show clearly that open reading frame 25
is a cob gene; this cob gene is designated cobN.

The SC510 Rif' Tn5Sp mutants 23, 13 and 12, which
have the Cob- phenotype, are mapped in open reading frame
24. These mutants are not complemented by plasmid pXL623,
which contains only the cobP gene. In contrast, these
mutants are complemented by plasmid pXL593 which contains
cobP and open reading frame 24, thereby indicating that
open reading frame 24 is responsible for their
complementation. Open reading frame 24 is hence a cob gene,
which is designated cobW.

EXAMPLE 5 - Genes and proteins

5.1 - 5.4-kb fragment

Five genes (cobA, cobB, cobC, cobD and cobE) are
hence defined on the 5.4-kb ClaI-HindIII-HindIII-HindIII
fragment. They code, respectively, for the following COB
proteins: COBA, COBB, COBC, COBD and COBE. The coding
portions of the genes (cobA to cobE) are described in
Figure 15, as well as the sequences of the COBA to COBE
proteins. Properties of each of these proteins are also
presented (amino acid composition, isoelectric point, index
of polarity and hydrophilicity profile).

5.2 - 8.7-kb fragment

Eight genes are hence defined on the 8.7-kb
fragment. These cobF to cobM genes code, respectively, for
the following COB proteins: COBF, COBG, COBH, COBI, COBJ,
COBK, COBL, and COBM. The coding portions of the genes
(cobF to cobM) are described in Figure 16, as well as the
sequences of the COBF to COBM proteins. Properties of each
of these proteins are also presented (amino acid
composition, molecular weight, isoelectric point, index of
polarity and hydrophilicity profile).

5.3 - 4.8-kb fragment

Three genes (cobX, cobS, cobT) are defined on the
4.8-kb SalI-SalI-SalI-SalI-SalI-BglI fragment. They code,
respectively, for the following proteins: COBX, COBS
and COBT. The coding portions of these genes are described
in Figure 40, as well as the sequences of the COBX, COBS

and COBT proteins. Arbitrarily, the ATG at position 1512 of
cobS has been chosen as the initiation codon, rather than
that located at position 1485 (see Figure 32). Properties
of each of these proteins are also shown (amino acid

5   composition, isoelectric point, index of polarity and
hydrophobicity profile). COBT possesses a hydrophilic
pocket corresponding to amino acids 214 to 305.

       5.4 - 3.9-kb fragment

       Two genes (cobU and cobV) are defined on the 3.9-

10   kb SstI-SstI-BamHI fragment. They code, respectively, for
the following proteins: COBU and COBV. The coding portions
of these genes are described in Figure 41, as well as the
sequences of the COBU to COBV proteins. Properties of each
of these proteins is also shown (amino acid composition,

15   isoelectric point, index of polarity and hydrophobicity
profile).

       5.5 - 13.4-kb fragment

       Five cob genes are defined on the 13.4-kb fragment
(cobQ, cobP, cobW, cobN and cobO and cobV). They code,

20   respectively, for the following proteins: COBQ, COBP, COBW,
COBN and COBO. The coding portions of these genes (cobQ,
cobP, cobW, cobN and cobO) are described in Figure 46, as
well as the sequences of COBQ, COBP, COBW, COBN and COBO
proteins. Properties of each of these proteins are also

25   shown (amino acid composition, isoelectric point, index of
polarity and hydrophobicity profile).

From the hydrophilicity profiles, which were produced according to the programmes of Hopp and Woods (1981), all the COB proteins with the exception of COBV are probably soluble proteins, as opposed to membrane proteins,

5 since the absence of large hydrophobic domains is noted. COBV is either a membrane protein, since 4 long hydrophobic domains are noted (see Figure 41), or a cytoplasmic protein having large hydrophobic domains.

For all the amino acid sequences of the COB

10 proteins, a methionine is indicated as the first amino acid at the NH_2-terminal position. It is understood that this methionine may be excised _in vivo_ (Ben Bassat and Bauer, 1984). Rules relating to the _in vivo_ excision of NH_2-terminal methionine by methionine aminopeptidase are known

15 to have been proposed. (Hirel et al., 1989).

Moreover, these protein sequences were compared with the Genpro proteins, Genpro being a Genbank protein extraction (version 59) augmented by putative coding portions larger than 200 amino acids, according to the

20 programme of Kanehisa (1984). No significant homology could be demonstrated with the parameters used on Genbank version 59, except for COBT. In effect, the COBT protein possesses a "core of acidic amino acids" between (amino acid) positions 224 and 293 (see Figure 40); in this portion of

25 the protein, more than one amino acid out of 2 is a glutamic or aspartic acid residue; this core of acidic amino acids renders the protein homologous over this

region, according to the programme of Kanehisa (1984), with other proteins also having such an acidic core. The most homologous proteins are: GARP protein of Plasmodium falciparum (Triglia et al., 1988), rat cardiac troponin T (Jin and Lin, 1989), human and rat prothymosin (Eschenfeld and Berger, 1986), an androgen-dependant rat protein that binds to spermine (Chang et al., 1987), and the human, rat and chicken "mid-size neurofilament subunit", proteins (Myers et al., 1987, Levy et al., 1987, Zopf et al., 1987). The function of these cores rich in acidic residues is unknown; however, this acidic core should either permit the binding of metal cations such as Co^{++}, which would give the COBT protein the role of a cobalt metallothionein, or else permit interactions with other proteins.

EXAMPLE 6 - Enzymatic studies

6.1 - Identification of COB proteins and their genes from purified enzymatic activities

This example describes how, from a purified protein, after its NH_2-terminal sequence has been established, it is possible to find the corresponding structural gene among sequenced cob genes.

6.1.1. Identification of the COBA protein encoded by the cobA gene

The purification of Pseudomonas denitrificans SUMT has been described (F. Blanche et al., 1989). The NH_2-terminal sequence of the protein thus purified could be

determined according to the technique described above. The first ten amino acids were identified:

1	2	3	4	5	6	7	8	9	10
Met	Ile	Asp	Asp	Leu	Phe	Ala	Gly	Leu	Pro

5   The NH_2 terminal sequence of the COBA protein (Figure 15) corresponds exactly to this sequence. The molecular weight of the purified SUMT, estimated by 12.5 % SDS-PAGE electrophoresis, is 30,000. The COBA protein has a molecular weight deduced from its sequence of 29,234

10   (Figure 15). The correspondences between the NH_2-terminal sequences and the molecular weights indicate clearly that the COBA protein corresponds to SUMT. The cobA gene is the SUMT structural gene.

         6.1.2. Identification of the COBB protein

15   encoded by the cobB gene

         a)   Assay of cobyrinic acid a,c-diamide synthase activity

         This example illustrates the assay of an activity of the pathway of biosynthesis of corrinoids which has never yet been described. The enzyme in question is

20   cobyrinic acid a,c-diamide synthase (CADAS), which catalyses the amidation of two carboxylic acid functions of the corrin or decobalt-ocorrin ring-system at positions a and c (Figure 17). The donor of the NH_2 group is L-glutamine, and the reaction consumes 1 molecule of ATP per

25   amidation of each carboxylic acid function. The assay which is described below applies to the diamidation reaction of cobyrinic acid; with a few modifications (detection in HPLC

at 330 nm in particular), it applies to the diamidation

reaction of hydrogenobyrinic acid.

The incubation mixture (0.1 M Tris-HCl pH 8

(250 μl)) containing ATP (1 mM), MgCl$_2$ (2.5 mM), glut-amine

5   (100 μm), cobyrinic acid (50 μM) or hydrogeno-byrinic acid

(50 μM) and cobyrinic a,c-diamide synthase (approximately 1

unit of activity) is incubated for 1 hour at 30°C. At the

end of the incubation, an aqueous solution (125 μl) of KCN

(2.6 g/l) and 0.2 M HCl (125 μl) are added to the mixture,

10   which is then heated to 80°C for 10 minutes and thereafter

centri-fuged for 5 minutes at 5,000 g. An aliquot (50 μl)

of the centrifugation supernatant is analysed in HPLC. It

is injected onto a 25-cm Nucleosil 5-C$_{18}$ column and eluted

with a gradient from 0 to 100 % of buffer B in A in the

15   course of 30 minutes; buffer A: 0.1 M potassium phosphate

pH 6.5, 10 mM KCN; buffer B: 0.1 M potassium phosphate pH

8, 10 mM KCN/acetonitrile (1:1). The corr-inoids are

detected by means of their UV absorption at 371 nm. The

unit of enzymatic activity is defined as the quantity of

20   enzyme necessary for synthesising 1 nmol of amide groups

per hour under the conditions described.

b)   Purification of Pseudomonas denitrificans cobyrinic

acid a,c-diamide synthase activity

This experiment illustrates how a Pseudomonas

25   denitrificans protein participating in the pathway of

biosynthesis of cobalamins may be purified.

Using the assay described in Example 6.1.2 a),
the purification of <u>Pseudomonas denitrificans</u> cobyrinic
acid a,c-diamide synthase is carried out as described
below.

5 In a typical purification experiment, wet cells
(7 g) of strain SC 510 Rif', into which plasmid pXL1500 has
been introduced (see Example 4.1. for the description of
pXL1500, as well as Figure 12), are suspended in 0.1 M
Tris-HCl pH 7.7 (30 ml) and sonicated for 15 minutes at

10 4°C. The crude extract is then recovered by centrifugation
for 1 hour at 50,000 g, and a portion (10 ml) of this
extract is injected onto a Mono Q HR 10/10 column
equilibrated with the same buffer. The proteins are eluted
with a linear KCl gradient (0 to 0.5 M). The fractions

15 containing the enzymatic activity (demonstrated by means of
the test described in Example 6.2 b)) are combined and
concentrated to 2.5 ml. After dilution with 25 mM Tris-HCl
pH 7.7 (1 ml), the proteins are fractionated on a Mono Q HR
5/5 using the above KCl gradient (0 to 0.5 M). The active

20 fractions are combined, and 0.1 M Tris-HCl pH 7.7 (1 ml)
containing 1.7 M ammonium sulphate is added to the sample,
which is then chromatographed on a Phenyl-Superose
(Pharmacia) column with a decreasing ammonium sulphate
gradient (1.0 M to 0 M). The fractions containing the

25 desired activity are combined and chromatographed on a
Bio-Gel HPHT (Bio-Rad) column with a potassium phosphate
gradient (0 to 0.35 M).

After this step, the enzyme is more than 95 %
pure. It shows no contaminant protein in SDS-PAGE. The
purity of the protein is confirmed by the uniqueness of the
NH$_2$-terminal sequence. Its molecular weight in this
technique is 45,000. The different steps of purifi-cation
of CADAS, with their purification factor and their yield,
are given in the table below.

Table: Purification of CADAS

Purification step	Vol (ml)	Proteins (mg)	Sp.activity (u/mg of proteins)	Yield	Purifi-cation factor[1]
Crude extract	10	200	8.5	-	-
MonoQ 10/10	12	15.1	108	96	12.7
MonoQ 5/5	3	3.75	272	60	32
Phenyl-Superose	1	0.865	850	43	100
Bio-Gel HPHT	2	0.451	1320	35	155

1 This factor is calculated from the increase in the specific activity
of the fractions during the purification.

c) NH$_2$-terminal sequence of Pseudomonas denitrificans
cobyrinic acid a,c-diamide synthase and identification of
the Pseudomonas denitrificans structural gene coding for
this activity

This example illustrates how the NH$_2$-terminal
sequence of a protein which participates in the pathway of
biosynthesis of cobalamins enables the structural gene
which codes for this protein to be identified.

The NH$_2$-terminal sequence of <u>Pseudomonas</u>
<u>denitrificans</u> cobyrinic acid <u>a,c</u>-diamide synthase, purified
as described in Example 6.1.2 b), was determined as
described above. 15 residues were identified:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ser	Gly	Leu	Leu	Ile	Ala	Ala	Pro	Ala	Ser	Gly	Ser	Gly	Lys	Thr

The NH$_2$-terminal sequence of the COBB protein
(Figure 15) corresponds exactly to this sequence except
that, in the sequence presented in Figure 15, a
methionine precedes the peptide sequence determined by
direct sequencing. It follows from this that the amino-
terminal methionine is definitely excised <u>in vivo</u> by
methionine aminopeptidase (Ben Bassat and Bauer, 1987).
The molecular weight of the purified CADAS, estimated
by 12.5 % SDS-PAGE electrophoresis, is 45,000. The COBB
protein has a molecular weight deduced from its
sequence of 45,676 (Figure 15). The correspondences
between the NH$_2$-terminal sequences and the molecular
weights indicate clearly that the COBB protein
corresponds to CADAS. The <u>cobB</u> gene is the CADAS
structural gene.

6.1.3. Identification of the COBI protein
encoded by the cobI gene

a)   Assay of an S-adenosyl-L-methionine:precorrin-2
methyltransferase activity

This example illustrates the assay of an
enzymatic activity of the pathway of biosynthesis of

corrinoids which has never yet been described. The enzyme in question is S-adenosyl-L-methionine: precorrin-2 methyltransferase (SP$_2$MT), which catalyses the transfer of a methyl group from S-adenosyl-L-methionine (SAM) to precorrin-2 to give precorrin-3 (Figure 18). Factors II and III, oxidation products of precorrin-2 and precorrin-3, respectively, have already been purified from cell extracts of Propionibacterium shermanii (Battersby and MacDonald, 1982, Scott et al., 1984); precorrin-2 and precorrin-3 are recognised as presumed intermediates of coenzyme B$_{12}$ biosynthesis, but they have never been purified as such. For this reason, the corresponding activity has never been either assayed or purified beforehand. The substrate of the enzymatic reaction, precorrin-2, is a very labile molecule which it is not possible to store, since it oxidises spontaneously in the presence of even infinitesimal traces of oxygen (Battersby and MacDonald, 1982). The principle of this enzymatic test hence lies in the possibility of generating precorrin-2 from SAM and δ-aminolevulinic acid at the required moment using an enzymatic extract of strain SC510 Rif^r into which plasmid pXL1500 has been introduced. The incubation must be performed under strictly anaerobic conditions.

The fractions containing SP$_2$MT are incubated in 0.1 M Tris-HCl pH 7.7 (1 ml) in the presence of

5 mM DTT, 1 mM EDTA, 100 μM [methyl-^{3}H]SAM (1 μCi),

0.8 mM δ-aminolevulinic acid and crude enzyme extract

(6 mg) of Pseudomonas denitrificans strain SC510 Rif'

pXL1500 for 3 hours at 30°C. Strain SC510 Rif' pXL1500

5   contains a strong SUMT activity (F. Blanche et al.,

1989). The tetrapyrrole compounds produced during the

incubation are bound to a DEAE-Sephadex anion exchange

column and esterified in methanol containing 5 % of

sulphuric acid in the absence of oxygen. The

10   dimethylated and trimethylated derivatives of uro'gen

III are then separated by thin-layer chromatography on

silica using dichloromethane/methanol (98.3:1.7) as an

eluent system (F. Blanche et al., 1989). The SP$_2$MT

activity is expressed as the ratio of the quantity of

15   trimethylated derivatives obtained to the total of (di-

and tri-)methylated derivatives produced, referred to

the quantity of protein. The SC510 Rif' pXL1500 extract

introduced in the test does not display detectable SP$_2$MT

activity under the assay conditions (the ratio of

20   precorrin-3 produced to precorrin-2 produced during the

test is less than 0.05).

b)   Purification of Pseudomonas denitrificans

S-adenosyl-L-methionine:precorrin-2 methyltransferase

       This experiment illustrates how a Pseudomonas

25   denitrificans protein participating in the pathway of

biosynthesis of cobalamins may be purified when an

assay for the activity in question exists.

The protein is purified from SC510 Rif' cells
containing plasmid pXL253. This is plasmid pKT230 into
which the 8.7-kb EcoRI fragment has been inserted
(Figure 13). In a typical purification experiment, wet
cells (50 g) of strain SC150 Rif' into which plasmid
pXL253 has been introduced are suspended in 0.1 M
potassium phosphate pH 7.7, 5 mM DTT (250 ml) and
sonicated for 15 minutes at 4°C. After centrifugation
at 50,000 g for 1 hour, the supernatant is passed
through a DEAE-Sephadex column (10 ml of gel) to remove
the tetrapyrrole compounds. The pH of the crude extract
thereby obtained is adjusted to pH 7.7 with 0.1 M KOH.
The proteins precipitating at between 33 % and 45 %
ammonium sulphate saturation are collected and
dissolved in 0.1 M Tris-HCl pH 7.7, 5 mM DTT (40 ml).
This solution is passed through a Sephadex G-25 column
eluted with 10 mM Tris-HCl pH 7.7, 5 mM DTT, and the
proteins collected are injected onto a DEAE-Trisacryl-M
column. The proteins are eluted with a linear gradient
of 0 to 0.25 M KCl, and the fractions containing the
SP$_2$MT activity are combined and passed a second time
through a Sephadex G-25 column as above. The protein
fraction is injected onto an Ultrogel HA (IBF) column
equilibrated in 10 mM Tris-HCl pH 7.7, 5 mM DTT. The
proteins are eluted with a linear gradient of 0 to
50 mM potassium phosphate pH 7.8 containing 5 mM DTT.
The fractions containing the desired activity are

combined and injected onto a MonoQ HR 5/5 (Pharmacia)
column equilibrated with 50 mM Tris-HCl pH 7.7,
5 mM DTT. The SP_2MT is eluted with a linear gradient (0
to 0.25 M) of KCl. At emergence from the MonoQ step,

5   12.5 % SDS-PAGE electrophoresis with staining with
silver salts reveals the enzyme is more than 99 % pure.
This is confirmed by the uniqueness of the NH_2-terminal
sequence of the protein. The molecular weight
calculated from the electrophoresis under denaturing

10   conditions (12.5 % SDS-PAGE) is 26,500. The steps of
purification of SP_2MT with their yields are described in
the table below.

Table: Purification of SP_2MT

15

Purification step	Vol (ml)	Proteins (mg)	Purification factor[1]
Crude extract	300	6000	-
Precipitation (33-45 %)	40	1530	3.9
DEAE-Tris-acryl-M	57	355	16.9
Ultrogel HA	30	71	85
MonoQ HR 5/5	12	33.5	179

30

[1] This factor is calculated from the yield of protein.

c) NH_2-terminal sequence of SP_2MT and identification of

35   the structural gene coding for this activity

This example illustrates how the NH_2-terminal sequence of a protein participating in the biosynthetic pathway enables the structural gene which codes for this protein to be identified. In the present example, the structural gene in question is that for SP_2MT.

The NH_2-terminal sequence of the purified protein was determined, as described above. The first 15 amino acids were identified:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ser	Gly	Val	Gly	Val	Gly	Arg	Leu	Ile	Gly	Val	Gly	Thr	Gly	Pro

The NH_2-terminal sequence of the COBI protein (Figure 16) corresponds exactly to this sequence except that, in the sequence presented in Figure 16, a methionine precedes the peptide sequence deduced from the nucleotide sequence. It follows from this that the amino-terminal methionine is definitely excised <u>in vivo</u> by methionine aminopeptidase (Ben Bassat and Bauer, 1987). The molecular weight of the purified SP_2MT, estimated by 12.5 % SDS-PAGE electrophoresis, is 26,500. The COBI protein has a molecular weight deduced from its amino acid sequence of 25,878 (Figure 16). The correspondences between the NH_2-terminal sequences and the molecular weights indicate clearly that the COBI protein corresponds to SP_2MT. The <u>cobI</u> gene is the SP_2MT structural gene.

6.1.4. Identification of the COBH protein encoded by the <u>cobH</u> gene

a)   Assay of precorrin-8x mutase activity

This example illustrates the assay of an enzymatic activity of the pathway of biosynthesis of cobalamines which has never been described hitherto. The enzyme in question is precorrin-8x mutase. This enzyme catalyses the transfer of the methyl group from position C-11 to position C-12 during the conversion of precorrin-8x to hydrogenobyrinic acid (see the nomenclature of the carbon atoms in Fig. 19; PL 68). More generally, it is the enzyme which catalyses the transfer of the methyl group C-11 to C-12, thereby leading to the corrin ring-system. The enzyme is referred to here as a mutase, although it has not been formally demonstrated that the transfer of the methyl group is intramolecular, even though this is very probable.

The enzymatic activity is demonstrated by the conversion of precorrin-8x (5 μM) to hydrogenobyrinic acid during incubations in the presence of enzyme fractions in 0.1 M Tris-HCl pH 7.7, 1 mM EDTA, at 30°C for 1 h. At the end of the incubation, the reaction is stopped by heating to 80°C for 10 min and, after centrifugation at 3000 × g for 10 min, the hydrogenobyrinic acid formed, present in the supernatant, is analysed by HPLC (see Example 6.1.2.a).

b)   Purification of precorrin-8x mutase.

The purification of <u>Pseudomonas denitrificans</u>
precorrin-8x mutase is carried out as described below.

During this purification, all the buffer
solutions are adjusted to pH 7.7.

5   In a typical purification experiment, cells
(50 g) of strain SC510 Rif', carrying plasmid pXL253
(plasmid pKT230 into which the 8.7-kb fragment has been
cloned at the <u>Eco</u>RI site, Figure 13) and obtained after
culture in PS4 medium, are resuspended in 0.1 M

10   potassium phosphate buffer (200 ml) and sonicated for
12 minutes. After centrifugation at 50,000 g for 1
hour, the supernatant is passed through a DEAE-Sephadex
column (10 ml of gel) to remove the tetrapyrrole
compounds. The pH of the solution is immediately

15   adjusted to 7.7 with 1 M KOH solution. The protein
fraction precipitating at between 40 and 60 % ammonium
sulphate saturation is collected by centrifugation and
dissolved in 0.1 M Tris-HCl (50 ml). This sample is
then injected onto an Ultrogel AcA 54 (IBF, France)

20   column (gel volume 1,000 ml) and the proteins are
eluted at a flow rate of 60 ml/h with 50 mM Tris-HCl.
The fractions containing the activity are pooled and
injected onto a DEAE-Trisacryl M (IBF, France) column
equilibrated with 50 mM Tris-HCl, and the proteins are

25   eluted with a gradient of 0 to 0.2 M KCl. The fractions
containing the protein to be purified are pooled and
passed through a Sephadex G-25 column equilibrated in

10 mM Tris-HCl. The protein fraction is injected onto
an Ultrogel HA (IBF, France) column equilibrated with
10 mM Tris-HCl, the proteins are eluted with a gradient
of 0 to 0.1 M potassium phosphate, and the active

5   fraction is then chromatographed on a Phenyl-Sepharose
CL (Pharmacia) 4B column in 10 mM potassium phosphate,
the column being eluted with a gradient of 0.65 to 0 M
ammonium sulphate. The active fractions are pooled. The
protein thereby obtained is more than 95 % pure

10   (according to the results of 12.5 % SDS-PAGE
electrophoresis and staining with silver salts). The
purity of the protein is confirmed by the uniqueness of
the N-terminal sequence. Its molecular weight
calculated using this technique is 22,000. The steps of

15   purification of precorrin-8x mutase with their
purification yields are described in the table below.

Table: Purification of precorrin-8x mutase

Purification step	Vol (ml)	Proteins (mg)	Purification factor[1]
Crude extract	250	6000	-
Precipitation (40-60 %)	50	2350	2.6
Ultrogel ACA 54	70	655	9.2
DEAE-Tris-acryl-M	30	271	22
Ultrogel HA	22	93	65
Phenyl-Sepharose	12	31	194

[1] This factor is calculated from the yield of protein.

c) NH$_2$-terminal sequence of precorrin-8x mutase and

identification of its structural gene

This example illustrates how the NH$_2$-terminal

sequence of a protein participating in the biosynthetic

pathway enables the structural gene which codes for

this protein to be identified.

The NH$_2$-terminal sequence of this protein was

determined as described above. 15 residues were

identified:

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Pro	Glu	Tyr	Asp	Tyr	Ile	Arg	Asp	Gly	Asn	Ala	Ile	Tyr	Glu	Arg

The NH$_2$-terminal sequence of the COBH protein

(Figure 16) corresponds exactly to this sequence except

that, in the sequence presented in Figure 16, a

methionine precedes the peptide sequence determined by

the sequencing described above. It follows from this

that the amino-terminal methionine is definitely

excised _in vivo_ by methionine aminopeptidase (Ben

Bassat and Bauer, 1987). Since the second residue is a

proline, this excision is in keeping with the rules

already stated (Hirel et _al_., 1989). The molecular

weight of the purified precorrin-8x mutase, estimated

by 12.5 % SDS-PAGE electrophoresis, is 22,000. The COBH

protein has a molecular weight deduced from its

sequence of 22,050 (Figure 16). The correspondences

between the NH$_2$-terminal sequences and the molecular

weights of these proteins indicate clearly that the

COBH protein corresponds to precorrin-8x mutase. cobH is the precorrin-8x mutase structural gene.

d)   Preparation, isolation and identification of precorrin-8x.

In a typical experiment for preparation of precorrin-8x, a crude enzyme extract of strain SC510 Rif' pXL253 (1000 mg of proteins) is incubated anaerobically for 20 h at 30°C in 0.1 M Tris-HCl buffer pH 7.7 (100 ml) with trimethylisobacteriochlorin (1000 nmol) prepared as described previously (Battersby et al., 1982), EDTA (1 mM), ATP (100 μmol), MgCl$_2$ (250 μmol), NADH (50 μmol), NADPH (50 μmol), SAM (50 μmol) and hydrogenobyrinic acid (20 μmol). At the end of the incubation, precorrin-8x is the preponderant tetrapyrrole product formed. It is isolated and purified by HPLC on a μBondapak C18 (Waters) column using a linear elution gradient of 0 to 50 % of acetonitrile in a potassium phosphate buffer pH 5.8. The mass of precorrin-8x (m/z = 880) and the mass of its methyl ester derivative (m/z = 978) indicate that it is a compound having the same empirical formula as hydrogenobyrinic acid. The UV/visible and fluorescence characteristics are very different from those of hydrogenobyrinic acid, and indicate that the molecule possesses two separate chromophors. Since the only enzymatic isomerisation reaction between precorrin-6x (Thibaut et al., 1990) and hydrogenobyrinic acid is the

migration of the methyl from C-11 to C-12, precorrin-8x

is the last intermediate before hydrogenobyrinic acid,

and the corresponding reaction is the migration of the

methyl from C-11 to C-12, catalysed by precorrin-8x

5   mutase.

6.1.5. Identification of the COBU protein

encoded by the cobU gene

a)   Assay of nicotinate-

nucleotide:dimethylbenzimidazole

10   phosphoribosyltransferase activity (Figure 5, reaction

5). This example illustrates the assay of an enzymatic

activity directly linked to the pathway of biosynthesis

of cobalamins. The enzyme in question is nicotinate-

nucleotide:dimethylbenzimidazole phosphoribosyl-

15   transferase (NN:DMBI PRT) (EC 2.4.2.21). The fractions

containing NN:DMBI PRT activity (approximately 5 units)

are incubated at 30°C for 8 min in 0.1 M glycine-NaOH

buffer pH 9.7 (500 μl) in the presence of 1 mM NaMN

(nicotinic acid mononucleotide) and 10 μM DMBI. The

20   reaction is then stopped by heating to 80°C for 10 min,

the reaction mixture is diluted with water (4 volumes)

and this solution (100 μl) is injected onto a 15-cm

Nucleosil 5-C8 HPLC column eluted with a 0.1 M

potassium phosphate pH 2.9/acetonitrile (93:7) mixture

25   at a flow rate of 1 ml/min. The α-ribazole 5'-phosphate

is detected and quantified by fluorimetry

(excitation : 260 nm; emission >370 nm). The unit of

enzymatic activity is defined as the quantity of enzyme necessary for generating 1 nmol of α-ribazole 5'-phosphate per hour under these conditions.

b) Purification of Pseudomonas denitrificans NN:DMBI PRT activity. This experiment illustrates how a P. denitrificans protein participating in the pathway of biosynthesis of cobalamins may be purified. Using the assay described in Example 6.1.5.a), the purification of Pseudomonas denitrificans NN:DMBI PRT is carried out as described below. In a typical purification experiment, wet cells (10 g) of strain SC510 Rif', into which plasmid pXL1490B has been introduced as described above, are used. Plasmid pXL1490B is described in Figure 38; this plasmid was obtained by cloning the 3.85-kb BamHI-SstI-SstI fragment of pXL519 (see Figure 38). This plasmid hence carries the cobU and cobV genes of P. denitrificans. The cells, cultured in PS4 medium supplemented with lividomycin, as described previously, are harvested after 96 hours of culture in PS4 medium. They are resuspended in 0.1M Tris-HCl buffer pH 7.2 (25 ml) and sonicated for 15 min at 4°C. The crude extract is then recovered by centrifugation for 1 h at 50,000 g, and thereafter passed through a DEAE-Trisacryl M (IBF, France) column equilibrated with the same buffer. 10% of the eluate (120 mg of proteins) is fractionated on a mono Q HR 10/10 column using a KCl gradient (0 to 0.6 M). The active fractions are pooled

and concentrated to 2 ml by ultrafiltration, and then, after mixing with 30 mM Tris-HCl buffer pH 7.2 (one volume), the sample is fractionated a second time on a Mono Q HR 5/5 column as before. The active fractions

5   are pooled, and the sample is then brought to a molarity of 1 M using ammonium sulphate and chromatographed on a Phenyl-Superose HR 5/5 column eluted with a decreasing ammonium sulphate gradient (1 M to 0 M). The fractions containing the desired

10   activity are pooled, concentrated by ultrafiltration and chromatographed on a Bio-Sil 250 gel permeation column eluted with 20 mM sodium phosphate/50 mM sodium sulphate pH 6.8.

After this step, the enzyme is more than 95%

15   pure. It shows no contaminant protein in SDS-PAGE. This purity is confirmed by the uniqueness of the NH_2-terminal sequence. Its molecular weight in this technique is 35,000. The different steps of purification of the NN:DMBI PRT are given in the table

20   below.

Table: Purification of P. denitrificans NN:DMBI PRT

Purification Step	Vol (ml)	Proteins (mg)	Sp. activity (u/mg of proteins	Yield	Purification factor'
Crude extract	6.0	120	2650	-	-
MonoQ 10/10	6.0	12.7	13515	51.3	5.1
MonoQ 5/5	3.0	6.19	20140	39.2	7.6
Phenyl-Superose	1.5	2.60	35510	29.0	13.4
Bio-Sil 250	1.2	1.92	39750	24.0	15.0

c) NH_2-terminal sequence of P. denitrificans NN:DMBI PRT
and identification of the Pseudomonas denitrificans
structural gene coding for this activity. The NH_2-
terminal sequence of Pseudomonas denitrificans NN:DMBI
PRT, purified as described in Example 6.1.5b), was
carried out according to the technique described above.
The first 15 residues were identified:

```
 1   2   3   4   5   6   7   8   9   10  11  12  13  14  15
Ser Ala Ser Gly Leu Pro Phe Asp Asp Phe Arg Glu Leu Leu Arg
```

The NH_2-terminal sequence of the COBU protein
(Figure 41) corresponds to this sequence except that,
in the sequence presented in Figure 41, a methionine
precedes the first amino acid of the peptide sequence
determined by direct sequencing. It follows from this
that the amino-terminal methionine is definitely
excised in vivo by methionine aminopeptidase (Ben
Bassat and Bauer, 1987). The molecular weight of the
purified N-transglycosidase, estimated by 12.5 % SDS-
PAGE electrophoresis, is 35,000. The COBU protein has a
molecular weight deduced from its sequence of 34,642

(Figure 41). The correspondences between the NH$_2$-
terminal sequences and the molecular weights clearly
indicate that the COBU protein corresponds to NN:DMBI
PRT. The cobU gene is the NN:DMBI PRT structural gene.

5   d) Specificity of NN:DBI PRT for DBI. This example
illustrates how a study of the specificity of P.
denitrificans NN:DMBI PRT enables P. denitrificans to
be made to biosynthesise various cobamides, using the
catalytic properties of P. denitrificans NN:DMBI PRT to

10   perform the synthesis of the nucleotide base in
question.

The enzyme substrate for synthesising
cobalamines is 5,6-dimethylbenzimidazole. Benzimidazole
and 5-methylbenzimidazole, respectively, are substrates

15   for the reaction with reaction rates of 157 % and 92 %,
respectively, compared to the natural substrate (5,6-
dimethylbenzimidazole), the NaMN concentration being
fixed at 2 mM. The specificity of P. denitrificans
NN:DMBI PRT is hence low for substrates containing a

20   benzimidazole ring-system. It is hence possible to use
P. denitrificans strain SC510 Rif^r (Cameron et al.,
1989), and to culture it in PS4 medium in which 5,6-
dimethylbenzimidazole is replaced by benzimidazole or
5-dimethylbenzimidazole, respectively, in order to make

25   the bacterium synthesise Coα-(benzimidazolyl)-Coβ-
cyanocobamide or Coα-(5-methylbenzimidazolyl)-Coβ-

cyanocobamide, respectively. There is no doubt that other cobamides could be synthesised in this way.

6.1.6. Identification of the COBV protein encoded by the cobV gene

This example illustrates how the assay of an activity of the pathway of biosynthesis of coenzyme B_{12} in P. denitrificans, and then the partial purification of this activity, can enable the structural gene for this enzyme to be identified in P. denitrificans.

a) Assay of GDP-cobinamide:α-ribazole-5'-phosphate cobinamidephosphotransferase (or cobalamin-5'-phosphate synthase) activity. This example illustrates the assay of an activity directly linked to the pathway of biosynthesis of cobalamines. The enzyme in question is cobalamin-5'-phosphate synthase. The fractions containing the activity (approximately 5 to 10 units) are incubated in darkness at 30°C in 0.3 Tris-HCl buffer pH 9.0 (500 μl) in the presence of 1 mM EDTA, 12.5 mM $MgCl_2$, 50 μM α-ribazole 5'-phosphate and 20 μM GDP-cobinamide [in 5'-deoxy-5'-adenosyl (Ado) or coenzyme form]. After 15 min of incubation, 20 mM potassium cyanide (500 μl) is added and the solution is heated to 80°C for 10 min. After centrifugation to remove the precipitated matter, the vitamin B_{12} 5'-phosphate present in the supernatant is assayed as described in Example 9. One unit of cobalamin-5'-phosphate synthase is defined as the quantity of enzyme

necessary for generating 1 nmol of cobalamine 5'-phosphate per h under the conditions described above.

Ado-GDP-cobinamide is obtained by incubation of Ado-cobinamide phosphate (Blanche et al., 1989) with a SC510 Rif' pXL623 extract under the conditions of assay of cobinamidephosphate guanylyltransferase (see 6.1.11.b). The α-ribazole and α-ribazole 5'-phosphate are isolated from SC510 Rif' cultures and purified by HPLC under the assay conditions described in Example 6.1.5a).

b) Partial purification of cobalamin-5'-phosphate synthase

This experiment illustrates how a P. denitrificans enzymatic activity participating in the pathway of biosynthesis of cobalamines of P. denitrificans can be partially purified. Using the assay described above, the purification of cobalamin 5'-phosphate synthase is carried out. For this purpose, in a typical purification experiment, wet cells (10 g) of strain SC510 Rif', into which plasmid pXL1490B has been introduced as described previously, are used. Plasmid pXL1490B is described in Figure 38: this plasmid corresponds to the 3.85-kb SstI-SstI-BamHI fragment cloned into PKT230. This plasmid carries the P. denitrificans cobU and cobV genes. The presence of this plasmid in P. denitrificans SC510 Rif' leads to an amplification of the cobalamin-5'-phosphate synthase

activity by a factor of approximately 100; it is hence probable that the insert carried by plasmid pXL1490B contains the structural gene for this enzyme; hence this gene can be only cobU or cobV. The SC510 Rif' pXL1490B cells are obtained by culture in PS4 medium supplemented with lividomycin, as described above. The cells are centrifuged and then resuspended in 0.1 M Tris-HCl (pH 8.3)/1 mM EDTA buffer (buffer A) (25 ml) and sonicated for 15 min at 4°C. The crude extract is then recovered by centrifugation for 1 h at 50,000 g and passed through a Sephadex G-25 column equilibrated with buffer A. The protein fraction is recovered and injected in 300-μl fractions (7.5 mg of proteins) onto a Superose 12 HR 10/30 column eluted in buffer A. The excluded fraction is recovered, mixed with an equal volume of buffer A/1.0 M ammonium sulphate and chromatographed on a Phenyl-Superose HR 5/5 column. The proteins are eluted with a decreasing ammonium sulphate gradient (0.5 M to 0 M) in buffer A, followed by a plateau at 0 M ammonium sulphate with the object of eluting the cobalamin-5'-phosphate synthase activity. The partial purification of this enzyme is described in the table below, on the basis of 75 mg of proteins introduced at the start in the purification process.

Table: Partial purification of P. denitrificans

cobalamin-5'-phosphate synthase

Purification step	Vol (ml)	Proteins (mg)	Sp. activity (u/mg of proteins)	Yield	Purification factor[1]
Crude extract	3.0	75	325	-	-
Superose 12HR	50.0	2.9	6,810	81	21
Phenyl-Superose	4.5	0.35	17,850	26	55

c) Specificity of cobalamin-5'-phosphate synthase. The Km for (Ado)GDP-cobinamide is 0.9 μm. However, the enzyme possesses the same affinity and a virtually identical reaction rate for the (CN, aq) form of the substrate. The Km of the enzyme for α-ribazole 5'-phosphate is approximately 2.7 μM. In addition, the purest preparations of cobalamine-5'-phosphate synthase catalyse the reaction of Ado-GDP-cobinamide with α-ribazole to give coenzyme B_{12} and, under these conditions, no accumulation of cobalamin 5'-phosphate is observed. The Km of the enzyme for α-ribazole is 7.8 μM. Intracellular α-ribazole 5'-phosphate and α-ribazole concentrations of 30 and 700 μM, respectively, were measured by HPLC during the production of cobalamins by SC510 Rif' in PS4 medium under the culture conditions described in Example 6.1.5a). This shows that coenzyme B_{12} may be generated directly from Ado-GDP-cobinamide by cobalamin-5'-phosphate synthase

without the participation of a cobalamin 5'-phosphatase.

The absence of accumulation or the presence of traces of cobalamin 5'-phosphate in the P. dinitrificans SC510 Rif' cultures confirms that coenzyme B_{12} is produced by the direct reaction of Ado-GDP-cobinamide with α-ribazole in vivo.

This direct reaction has already been observed and described in vitro in Propionibacterium shermanii (Ronzio et al., 1967; Renz, 1968). As the cobalamin-5'-phosphate synthase structural gene can be only cobU or cobV, since the amplication in P. denitrificans of a fragment carrying these two P. denitrificans cob genes leads to an increase in cobalamin-5'-phosphate synthase activity by a factor of 100, and since the cobU gene is the NN:DMBI PRT structural gene, cobV is hence the cobalamin-5'-phosphate synthase structural gene.

6.1.7. Identification of the COBK protein encoded by the cobK gene

a) Assay of precorrin-6x reductase activity.

This example illustrates the assay of a novel enzymatic activity directly linked to the pathway biosynthesis of cobalamins. The enzyme in question is precorrin-6x reductase.

The fractions containing precorrin-6x reductase activity (approximately 0.05 unit, U) are

incubated at 30°C for 60 min in 0.1 M Tris-HCl buffer pH 7.7 (250 μl) in the presence of 1 mM EDTA, 500 μM NADPH, 25 μM [methyl-^{3}H]SAM (80 μCi/μmol), 4 μM precorrin-6x (Thibaut et al., 1990) and partially

5   purified dihydroprecorrin-6x methylase (0.5 U) (see preparation below). The reaction is then stopped by heating to 80°C for 5 min and, after centrifugation at 5000 × g for 5 min, the supernatant is injected onto a DEAE-Sephadex column (containing 200 μl of gel). The

10   column is then washed extensively with the Tris-HCl buffer, and the compounds bound are eluted with 1 M HCl (4 ml). The radio activity in this eluent is counted by liquid scintillation counting. The unit of enzymatic activity is defined as the quantity of enzyme necessary

15   for reducing 1 nmol of precorrin-6x per h under these conditions.

Dihydroprecorrin-6x methylase is partially purified from a crude extract of SC510 Rif' pXL253 on a Mono Q HR 5/5 (Pharmacia) anion exchange column. The

20   column is eluted with a linear gradient of 0 to 0.4 M KCl in 0.1 M Tris-HCl buffer pH 7.7. The enzymatic activity is eluted at 0.35 M KCl. This activity is detected and quantified by means of the precorrin-6x reductase activity test defined above (in the presence

25   of precorrin-6x reductase (0.5 U) in the incubation medium). After the Mono Q step, the fractions containing dihydroprecorrin-6x methylase activity are

completely devoid of precorrin-6x reductase activity.
The unit of methylase activity is defined as the
quantity of enzyme necessary for transferring 1 nmol of
methyl groups to dihydroprecorrin-6x per h under the
conditions described above.

b) Purification of precorrin-6x reductase activity

.Using the assay described above, the
purification of Pseudomonas denitrificans precorrin-6x
reductase is carried out as described below.

In a typical purification experiment, wet
cells (100 g) of strain SC510 Rif', into which plasmid
pXL253 (plasmid pKT230 into which the 8.7-kb fragment
has been cloned at the EcoRI site, Figure 13) has been
introduced, are suspended in 0.1 M Tris-HCl pH 7.7/1 mM
EDTA buffer (buffer A) (200 ml) and sonicated for 15
min at 4°C. The crude extract is then recovered by
centrifugation at 1 h at 50,000 × g and passed in three
portions through a Sephadex G-25 column equilibrated
with buffer A. The three fractions excluded from the
gel are pooled and adjusted to 1 l with buffer A. The
proteins precipitating at between 25 and 40 % ammonium
sulphate saturation are collected by centrifugation and
resuspended in buffer A (50 ml), and this solution is
desalted through a Sephadex G-25 column equilibrated
with buffer B (25 mM Tris-HCl/500 µM DTT/15% glycerol).
The protein solution is then injected at 2.5 ml/min
onto a Q Sepharose Fast Flow (Pharmacia) column

equilibrated with buffer B, and the proteins are eluted with a buffer B/0.2 M KCl mixture. This fraction is desalted on a Sephadex G-25 column equilibrated with buffer C (50 mM Tris-HCl/500 μM DTT/15% glycerol). The

5   protein solution is then fractionated (100 mg of proteins at each chromatographic run) on a Mono Q HR 10/10 (Pharmacia) column using a gradient of 0 to 0.4 M KCl in buffer C, and the fraction containing the activity is thereafter chromatographed on a Phenyl-

10 Superose HR 10/10 (Pharmacia) column in a linear decreasing ammonium sulphate gradient (1 to 0 M). The active fraction is desalted and the precorrin-6x reductase is repurified on a Mono Q HR 5/5 column. It is eluted in 50 mM Tris-HCl pH 8.1/500 μM DTT/15%

15 glycerol buffer with a gradient of 0 to 0.2 M KCl. To complete the purification, the protein is finally chromatographed on a Bio-Sil 250 (Bio-Rad) column eluted with 20 mM potassium phosphate/50 mM sodium sulphate pH 6.8/500 μM DTT/15% glycerol. After this

20 step, the enzyme is more than 95 % pure. It shows no contaminant protein in SDS-PAGE, the proteins being visualised with silver nitrate. This degree of purity is confirmed by the uniqueness of the NH_2-terminal sequence. Its molecular weight in this technique is

25 31,000. The different steps of purification of precorrin-6x reductase with their purification factor and their yield, are given in the table below.

Table: Purification of precorrin-6x reductase

Purification step	Vol (ml)	Proteins (mg)	Sp.activity (u/mg of proteins)	Yield	Purific- ation factor[1]
Crude extract	270	9600	0.535	-	-
A.S. 25 40%	100	4160	1.14	92	2.1
Q Sepharose	150	1044	3.64	74	6.8
Mono Q 10/10	55	67	24.5	32	46
Phenyl-Superose	10	2.2	325	14	607
Mono Q 5/5	2.5	0.082	5750	9.2	10750
Bio-sil 250	1.0	0.055	7650	8.2	14300

c) NH$_2$-terminal sequence and partial internal sequences of Pseudomonas denitrificans precorrin-6x reductase and identification of the Pseudomonas denitrificans structural gene coding for this activity

The NH$_2$-terminal sequence of Pseudomonas denitrificans precorrin-6x reductase, purified as described above, was determined as described before. Six residues were identified:

Ala-Gly-Ser-Leu-Phe-Asp

Similarly, after tryptic digestion and separation of the fragments by HPLC on a C-18 reversed-phase column, three internal sequences were obtained:

Ile-Gly-Gly-Phe-Gly-Gly-Ala-Asp-Gly-Leu

Arg-Pro-Glu-Trp-Val-Pro-Leu-Pro-Gly-Asp-Arg

Val-Phe-Leu-Ala-Ile-Gly

The NH$_2$-terminal sequence of the COBK protein (Figure 16) corresponds exactly to the NH$_2$-terminal sequence of precorrin-6x reductase except that, in the

sequence presented in Figure 16, a methionine precedes

the peptide sequence determined by direct sequencing.

It follows from this that the amino-terminal methionine

is definitely excised in vivo by methionine

5   aminopeptidase (Ben Bassat and Bauer, 1987). Similarly,

the three internal sequences correspond to the three

sequences 60 to 69, 112 to 122 and 143 to 148 of the

COBK protein. The molecular weight of the purified

precorrin-6x reductase is estimated by SDS-PAGE

10   electrophoresis at 31,000. The COBK protein has a

molecular weight deduced from its sequence of 28,000

(Figure 16). The correspondences between the internal

NH_2-terminal sequences and the molecular weights

indicate clearly that the COBK protein corresponds to

15   precorrin-6x reductase. The cobK gene is the precorrin-

6x reductase structural gene.

d) Reaction catalysed by precorrin-6x reductase

        The enzymatic reaction of reduction of

precorrin-6x is strictly NADPH-dependant in P.

20   denitrificans. NADPH cannot be replaced by NADH. When

the purified enzyme (or an active fraction during

purification, or even a crude enzyme extract) is

incubated under the conditions of the assay of

activity, but in the absence of SAM and of

25   dihydroprecorrin-6x methylase, the product of the

reaction can then be purified by HPLC in the system

described for the purification of precorrin-6x (see

Example 6.1.4.d). After desalting and esterification

(4% methanolic sulphuric acid, 20°C, 24 h, argon

atmosphere), the corresponding ester has a mass m/z =

1008. The product of the reaction catalysed by

5   precorrin-6x reductase is hence dihydroprecorrin-6x,

also known as precorrin-6y.

6.1.8. Identification of the COBQ protein

encoded by the cobQ gene

a) Assay of cobyric acid synthase activity

10   This example illustrates the assay of an

enzymatic activity of the pathway of biosynthesis of

cobalamins which has never been described hitherto. The

enzyme in question is cobyric acid synthase. This

enzyme catalyses the amidation of the peripheral

15   carboxylic acid functions at positions b, d, e and g on

the corrin ring-system (see Fig. 19; PL. 68). The NH_2-

group donor is L-glutamine, and each amidation reaction

is accompanied by the consumption of one ATP molecule.

The fraction to be assayed is incubated in

20   darkness at 30°C for 60 min in 0.1 M Tris hydrochloride

buffer pH 7.5 (250 μl) containing 1 mM DTT, 1 mM EDTA,

1 mM ATP, 2.5 mM $MgCl_2$, 1 mM glutamine and 10 μM Ado-

cobyrinic acid di- or pentaamide. The reaction is

stopped by adding 0.1 M aqueous potassium cyanide

25   solution (25 μl). After heating to 80°C for 10 min and

centrifugation at 3000 × g for 10 min, the compounds

formed, present in the supernatant, are analysed by

HPLC. The unit of activity is defined as the quantity of enzyme necessary for generating 1 nmol of amide functions per h under these conditions.

5'-Deoxy-5'-adenosyl(Ado)-cobyrinic acid diamide and pentaamide are isolated from cultures of strain SC510 in PS4 medium, using the method the principle of which is described in Example 9.

b) Purification of cobyric acid synthase

Using the assay described in Example 6.1.8 a), purification of Pseudomonas denitrificans cobyric acid synthase is carried out as described below.

In a typical purification experiment, wet SC510 Rif' cells (6 g), into which strain plasmid pXL618 (see Example 4.5.2) has been introduced, are sonicated in 0.1 M Tris-HCl pH 7.7, 1 mM DTT, 1 mM EDTA buffer (15 ml). After centrifugation (50,000 × g for 1 h), the extract is brought to 20 % of glycerol (vol/vol). 10 mM Tris-HCl, 1 mM DTT, 20% glycerol buffer (24 ml) are added to the crude extract (8.5 ml; 203.5 mg of proteins). The solution is injected onto Mono Q HR 10/10 (Pharmacia) at 2 ml/min, equilibrated with 50 mM Tris-HCl pH 7.7, 1 mM DTT, 20% glycerol buffer. The proteins are eluted with a linear gradient of 0.5 M NaCl and the active fractions are pooled and brought to 1 mM EDTA. The solution is brought to 0.85 M with respect to ammonium sulphate and injected onto a Phenyl-Superose HR 5/5 (Pharmacia) column equilibrated

in Tris-HCl pH 7.7, 1 mM DTT, 0.85 M ammonium sulphate

buffer, and the proteins are eluted with a linear

decreasing gradient of 0.85 M to 0 M ammonium sulphate.

The fractions are immediately brought to 20% of

5   glycerol. The active fraction is concentrated to 2.5 ml

by ultrafiltration and chromatographed on a PD 10

(Pharmacia) column equilibrated and eluted with 50 mM

Tris-HCl pH 8.3, 1 mM DTT, 20% glycerol (vol/vol)

buffer. The protein fraction is collected and injected

10   onto a mono Q HR 5/5 column equilibrated with the same

buffer, and the proteins are eluted with a linear

gradient of 0.5 M NaCl. Gel permeation chromatography

on Bio-Sil 250 (Bio-Rad) gel in 50 mM Tris-HCl pH 7.5,

1 mM DTT, 20% glycerol, 0.1 M NaCl buffer medium

15   finally enables a protein which is more than 97 % pure

to be obtained. It shows no contaminant protein in SDS-

PAGE. This purity is confirmed by the uniqueness of the

NH_2-terminal sequence. Its molecular weight in this

technique is 57,000. The different steps of

20   purification of cobyric acid synthase with their

purification factor and their yield are given in the

table below.

Table: Purification of cobyric acid synthase

Purification step	Vol (ml)	Proteins (mg)	Sp. activity U/mg a A	b B	Yield[c]	Purification factor[i]
Crude extract	8.5	203	114 /	118	-	-
Mono Q 10/10	8.0	35.5	388 /	425	60	3.4
Phenyl-Superose	8.0	3.23	1988 /	2021	28	17
Mono Q 5/5	1.0	1.20	4549 /	4085	24	40
Bio-Sil 250	0.75	0.88	4992 /	N.D.	19	44

a / with Ado-cobyrinic acid a,c-diamide as substrate

b / with Ado-cobyrinic acid pentaamide as substrate

ND = Not Determined

The very high degree of purity of the purified protein, together with the constancy of the ratio of the activities of amidation of cobyrinic acid diamide and pentaamide throughout the process of purification of the protein (see table above), indicate unambiguously that one and the same protein is responsible for the four activities of amidation of the corrin ring-system at positions b, d, e and g.

c) NH$_2$-terminal sequence of Pseudomonas denitrificans cobyric acid synthase and identification of the Pseudomonas denitrificans structural gene coding for this activity

The NH$_2$-terminal sequence of Pseudomonas denitrificans cobyric acid synthase was determined as described above. Sixteen residues were identified:

Thr-Arg-Arg-Ile-Met-Leu-Gln-Gly-Thr-Gly-Ser-Asp-Val-
Gly-Lys-Ser

The NH_2-terminal sequence of the COBQ protein
(Figure 47) corresponds exactly to this sequence except
that, in the sequence presented in Figure 47, a
methionine precedes the peptide sequence determined by
direct sequencing. It follows from this that the amino-
terminal methionine is definitely excised in vivo by
methionine aminopeptidase (Ben Bassat and Bauer, 1987).
The molecular weight of the purified cobyric acid
synthase is estimated by SDS-PAGE electrophoresis at
57,000. The COBQ protein has a molecular weight deduced
from its sequence of 52,000 (Figure 47). The
correspondences between the NH_2-terminal sequences and
the molecular weights indicate clearly that the COBQ
protein corresponds to cobyric acid synthase. The cobQ
gene is the cobyric acid synthase structural gene.

6.1.9. Identification of the COBO protein
encoded by the cobO gene

a) Assay of cob(I)alamin adenosyltransferase (EC
2.5.1.17) activity

This example illustrates the assay of an
enzymatic activity directly linked to the pathway of
biosynthesis of cobalamins. The enzyme in question is
cob(I)alamin adenosyltransferase (EC 2.5.1.17). This
enzyme was demonstrated in bacterial cells (Ohta et
al., 1976, Brady et al., 1962) and animal cells (Fenton

et al., 1978). It was purified from <u>Clostridium</u> <u>tetanomorphum</u> (Vitols et al., 1966).

The fractions containing cob(I)alamin adenosyltransferase activity (approximately 20 units) are incubated anaerobically at 30°C for 15 min protected from light in 0.2 M Tris-HCl buffer pH 8.0 (1 ml) in the presence of 5 mM DTT, 400 μM [8-^{14}C]-ATP (2.5 μCi/μmol), 800 μM MnCl$_2$, 50 μM hydroxocobalamin or diaquacobinamide and KBH$_4$ (3 mg). The reaction is then stopped by heating to 80°C for 10 min and, after centrifugation at 15000 × g for 5 min, the supernatant (200 μl) is analysed by HPLC (Gimsing et al., 1986, Jacobsen et al., 1986).

The unit of enzymatic activity is defined as the quantity of enzyme necessary for generating 1 nmol of adenosylcorrinoid per min under these conditions.

b) Purification of cob(I)alamin adenosyltransferase activity

Using the assay described in Example 6.1.9 a), the purification of <u>Pseudomonas denitrificans</u> cob(I)alamin adenosyltransferase is carried out as described below.

In a typical purification experiment, wet cells (10 g) of strain SC510 Rif^r in which the cobO gene has been amplified are suspended in 0.2 M Tris-HCl buffer pH 8.0 (20 ml) and sonicated for 40 min at 4°C. The crude extract is then recovered by centrifugation

for 1 h at 50,000 × g and desalted on PD10 (Pharmacia)

columns equilibrated with 50 mM Tris-HCl pH 8.0, 5 mM

DTT buffer (buffer A). The protein solution is then

fractionated (280 mg of proteins at each

5   chromatographic run) on a Mono Q HR 10/10 (Pharmacia)

column using a gradient of 0 to 0.5 M KCl in buffer A,

and the fractions containing the activity are then

pooled, concentrated by ultrafiltration and

chromatographed on a Phenyl-Superose HR 10/10

10   (Pharmacia) column in a linear decreasing ammonium

sulphate gradient (1.7 to 0 M), the column being

equilibrated in 0.1 M Tris-HCl pH 8.0, 5 mM DTT buffer.

To complete the purification, the protein is finally

chromatographed, after concentration by

15   ultrafiltration, on a Bio-Sil 250 (Bio-Rad) column

eluted with 50 mM Tris-HCl pH 7.5, 0.1 M NaCl, 5 mM DTT

buffer.

         After this step, the enzyme is more than 95 %

pure. It does not show any contaminant protein in SDS-

20   PAGE. Its molecular weight in this technique is 28,000.

This degree of purity is confirmed by the uniqueness of

the NH_2-terminal sequence. The different steps of

purification of cob(I)alamin adenosyltransferase, with

their purification factor and their yield, are given in

25   the table below for the following two substrates:

diaquacobinamide (a) and hydroxocobalamin (b). These

results demonstrate the absence of specificity of this

enzyme for the nature of the corrinoid substrate.
Moreover, all corrinoids of the biosynthetic pathway
between cobyrinic acid diamide and B_{12} have been
isolated (Blanche et al., unpublished results) in their
native form, and have proved to be in coenzyme form.
This demonstrates that the natural substrate of
cob(I)alamin adenosyltransferase is cobyrinic acid a,c-
diamide.

Table: Purification of cob(I)alamin adenosyltransferase

Purification step	Vol (ml)	Proteins (mg)	Sp. activity U/mg a A / b B		Yield[*]	Purific-ation factor[i]
Crude extract[c]	100	1400	5.4 /	3.4	-	-
Mono Q 10/10	90	140	34.9 /	14.1	65	6.5
Phenyl-Superose	30	15.9	84.5 /	49.5	18	16
Bio-Sil 250	6.5	2.9	182.4 /	88.7	7.0	34

c/ after desalting on PD10

c) NH_2-terminal sequence of Pseudomonas denitrificans
cob(I)alamin adenosyltransferase and identification of
the Pseudomonas denitrificans structural gene coding
for this activity.

The NH_2-terminal sequence of Pseudomonas
denitrificans cob(I)alamin adenosyltransferase,
purified as described in Example 6.1.9 b), was
determined as described above. 13 residues were
identified:

Ser-Asp-Glu-Thr-?-Val-Gly-Gly-Glu-Ala-Pro-Ala-Lys-Lys

The NH$_2$-terminal sequence of the COBO protein
(Figure 47) corresponds exactly to the NH$_2$-terminal
sequence of cob(I)alamin adenosyltransferase except
that, in the sequence presented in Figure 47, a

5   methionine precedes the peptide sequence determined by
direct sequencing. It follows from this that the amino-
terminal methionine is definitely excised in vivo by
methionine aminopeptidase (Ben Bassat and Bauer, 1987).
The molecular weight of the purified cob(I)alamin

10   adenosyltransferase is estimated by SDS-PAGE
electrophoresis at 28,000. The COBO protein has a
molecular weight deduced from its sequence of 24,000
(Figure 47). The correspondences between NH$_2$-terminal
sequences and the molecular weights indicate clearly

15   that the COBO protein corresponds to cob(I)alamin
adenosyltransferase. The cobO gene is the cob(I)alamin
adenosyltransferase structural gene.

6.1.10. Identification of the COBN protein
encoded by the cobN gene

20   a) Demonstration of the activity of conversion of
hydrogenobyrinic acid a,c-diamide to cobyrinic acid
a,c-diamide

This example illustrates the demonstration of
an enzymatic activity directly linked to the pathway of

25   biosynthesis of cobalamins which has never been
described hitherto. The activity in question is that of

conversion of hydrogenobyrinic acid a,c-diamide to cobyrinic acid a,c-diamide.

This activity is demonstrated, inter alia, by the following typical experiment. A crude extract of strain SC510 Rif' is obtained by sonication of wet cells (10 g) in 0.2 M Tris-HCl buffer pH 8.0 (20 ml), followed by removal of the cell debris by centrifugation for 1 h at 50,000 × g. Proteins (1000 mg) of this extract are incubated for 1 h at 30°C with carbon-14-labelled hydrogenobyrinic acid diamide (32 nmol; 50 μCi/μmol) in 0.2 M Tris-HCl buffer pH 8.0 (40 ml) containing 7 mM ATP and 200 μM $CoCl_2$. The reaction is stopped by adding 1 M KH_2PO_4 (7.5 ml) and 0.3 M KCN (6 ml), followed by heating for 10 min at 80°C. After centrifugation at 15000 × g for 50 min, HPLC analysis of the supernatant shows: (1) the formation during the incubation of cobyrinic acid a,c-diamide (19.2 nmol) having the same specific radioactivity as the starting hydrogenobyrinic acid a,c-diamide, and (2) the disappearance of a corresponding quantity of the latter. To confirm that the product is indeed cobyrinic acid a,c-diamide, the product is purified by HPLC and then esterified in methanol containing 5 % of sulphuric acid (18 h, 20°C). The authenticity of the cobyrinic acid a,c-diamide pentamethyl ester produced is demonstrated by TLC (relative to a reference sample) and mass spectrometry. It should be noted that, under

similar incubation conditions in which the radioactive labelling is introduced, not into the hydrogenobyrinic acid a,c-diamide, but into the cobalt (using cobalt-57), cobalt-57-labelled cobyrinic acid a,c-diamide is biosynthesised and the same conclusions could be drawn.

Carbon-14-labelled hydrogenobyrinic acid a,c-diamide is obtained in the following manner: hydrogenobyrinic acid is biosynthesised in vitro using [methyl-^{14}C]SAM, then converted to hydrogenobyrinic acid a,c-diamide and purified by HPLC as described in Example 6.1.2.

This study demonstrates that the insertion of cobalt takes place at hydrogenobyrinic acid a,c-diamide level in P. denitrificans. Under the conditions described, hydrogenobyrinic acid is not a substrate for enzymatic chelation with cobalt.

b) Assay and purification of a protein of strain SC510 Rif' involved in the conversion of hydrogenobyrinic acid a,c-diamide to cobyrinic acid a,c-diamide

The fraction to be assayed (0.5 to 2 units) is incubated for 60 min at 30°C with crude extract (50 μl) of strain SC510 Rif' obtained as described above, 7 mM ATP, 200 μM CoCl$_2$, and 7 μM carbon-14-labelled hydrogenobyrinic acid a,c-diamide (50 μCi/μmol) in 0.1 M Tris-HCl buffer pH 8.0 (400 μl). The reaction is stopped by adding 1 M KH$_2$PO$_4$ (75 μl) and 0.3 M KCN (60 μl), followed by heating for 10 min at 80°C. After

centrifugation at 15000 × g for 15 min, the supernatant
is analysed by HPLC in order to quantify the cobyrinic
acid a,c-diamide formed (see Example 9). The unit of
enzymatic activity is defined as the quantity of enzyme

5   necessary for generating 1 nmol of cobyrinic acid a,c-
diamide per h under these conditions. Under these
conditions, it is apparent that extracts of strain
SC510 Rif' into which plasmid pXL1909 has been
introduced (see Example 4.5.2) possess an activity

10   between 20 and 50 times as high as extracts of strain
SC510 Rif'. It is on this basis that a protein which is
alone responsible for this amplication of activity is
purified.

In a typical purification experiment, wet

15   cells (10 g) of strain SC510 Rif', into which plasmid
pXL1909 has been introduced, are suspended in 0.2 M
Tris-HCl buffer pH 8.0 (20 ml) and sonicated for 30 min
at 4°C. The crude extract is then recovered by
centrifugation for 1 h at 50,000 × g and desalted on

20   PD10 (Pharmacia) columns equilibrated with 0.1 M Tris-
HCl buffer pH 8.0 (buffer A). The protein solution is
then fractionated (213 mg of proteins at each
chromatographic run) on a Mono Q HR 10/10 (Pharmacia)
column using a gradient of 0 to 0.5 M KCl in buffer A,

25   and the fractions containing the activity are then
pooled, concentrated by ultrafiltration, desalted on
PD10 (Pharmacia) columns equilibrated with 0.1 M Tris-

HCl buffer pH 7.2 (buffer B) and chromatographed on a
Mono Q HR 10/10 (Pharmacia) column using a gradient of
0 to 0.5 M KCl in buffer B. The fractions containing
the activity are pooled, concentrated by
ultrafiltration, desalted on PD10 (Pharmacia) columns
equilibrated with buffer B and chromatographed on a
Mono Q HR 5/5 (Pharmacia) column using a gradient of 0
to 0.5 M KCl in buffer B. To complete the purification,
the protein is finally chromatographed on a Bio-Sil 250
(Bio-Rad) column eluted with 20 mM potassium
phosphate/50 mM sodium sulphate pH 6.8.

After this step, the enzyme is more than 95 %
pure. It does not show any contaminant protein in SDS-
PAGE. Its molecular weight in this technique is
135,000. This degree of purity is confirmed by the
uniqueness of the NH$_2$-terminal sequence. The different
steps of purification of the protein of strain SC510
Rif' involved in the conversion of hydrogenobyrinic acid
a,c-diamide to cobyrinic acid a,c-diamide, with their
purification factor and their yield, are given in the
table below.

Table: Purification of a protein of strain SC510 Rif'
involved in the conversion of hydrogenobyrinic acid
a,c-diamide cobyrinic acid a,c-diamide

Purification step	Vol (ml)	Proteins (mg)	Sp. activity (u/mg of proteins)	Yield	Purification factor'
Crude extract	31.5	1278	0.23	-	-
Mono Q 10/10	44	79.2	2.4	64	10
Mono Q 10/10	21	33.6	6.8	78	30
Mono Q 5/5	3	6.6	16.0	36	70
Bio-Sil 250	1.8	5.9	16.3	33	71

c) NH$_2$-terminal sequence of the Pseudomonas
denitrificans protein involved in the conversion of
hydrogenobyrinic acid a,c-diamide to cobyrinic acid
a,c-diamide, and identification of the Pseudomonas
denitrificans structural gene coding for this activity

The NH$_2$-terminal sequence of this protein,
purified as described in Example 6.1.10b), was
determined as described above. Six residues were
identified:

His-Leu-Leu-Leu-Ala-Gln

The NH$_2$-terminal sequence of the COBN protein
(Figure 47) corresponds exactly to the NH$_2$-terminal
sequence of the purified protein except that, in the
sequence presented in Figure 47, a methionine precedes
the peptide sequence determined by direct sequencing.
It follows from this that the amino-terminal methionine

is definitely excised in vivo by methionine
aminopeptidase (Ben Bassat and Bauer, 1987). The
molecular weight of the purified protein is estimated
by SDS-PAGE electrophoresis at 135,000. The COBN

5   protein has a molecular weight deduced from its
sequence of 138,000 (Figure 47). The correspondences
between the NH_2-terminal sequences and the molecular
weights indicated clearly that the COBN protein
corresponds to the protein involved in the conversion

10   of hydrogenobyrinic acid a,c-diamide to cobyrinic acid
a,c-diamide. The cobN gene is hence the structural gene
for this protein.

6.1.11. Identification of the COBP protein
encoded by the cobP gene

15   a) Assay of cobinamide kinase activity

This example illustrates the assay of an
enzymatic activity of the pathway of biosynthesis of
cobalamins which has never been studied hitherto. The
activity in question is that of cobinamide kinase. It

20   catalyses the ATP-dependent phosphorylation of the
hydroxyl group of the (R)-1-amino-2-propanol residue of
Ado-cobinamide to generate cobinamide phosphate.

The fraction to be assayed is incubated in
darkness at 30°C for 60 min in 0.1 M Tris-HCl buffer pH

25   8.8 (500 μl) containing 1 mM EDTA, 1 mM ATP, 2.5 mM
$MgCl_2$ 16 μM Ado-cobinamide (Blanche et al., 1989). The
reaction is stopped by adding 20 mM aqueous potassium

cyanide solution (500 μl). After heating to 80°C for 10 min and centrifugation at 5,000 × g for 10 min, the cobinamide phosphate formed, present in the supernatant, is assayed by HPLC (see Example 9) using the following simplified linear gradient: 25 % to 30 % of B in A in the course of 15 min, then 30 % to 100 % of B in the course of 12 min, and 3 min at 100 % of B.

The unit of activity is defined as the quantity of enzyme necessary for generating 1 nmol of cobinamide phosphate from cobinamide per h under these conditions.

b) Assay of cobinamidephosphate guanylyltransferase activity

This example illustrates the assay of an enzymatic activity of the pathway of biosynthesis of cobalamins which has never been studied hitherto. The activity in question is that of cobalamidephosphate guanylyltransferase. It catalyses the addition of the GMP portion of a GTP molecule to Ado-cobinamide phosphate, thereby generating one molecule of GDP-cobinamide and liberating one molecule of pyrophosphate.

This activity is assayed under the same conditions as cobinamide kinase, except that Ado-cobinamide phosphate (16 μM) (Blanche et al., 1989) and GTP (2 mM) replace Ado-cobinamide and ATP, respectively, during the incubation.

The unit of activity is defined as the
quantity of enzyme necessary for generating 1 nmol of
GDP-cobinamide from cobinamide phosphate per h under
these conditions.

5   c) Purification of cobinamide kinase

Using the assay described in Example
6.1.11a), the purification of _Pseudomonas denitrificans_
kinase is carried as described below.

In a typical purification experiment, wet
10   SC510 Rif' cells (5 g), into which strain plasmid pXL623
has been introduced (see Example 4.5.2) are sonicated
in 0.1 M Tris buffer pH 7.6 (buffer A) (20 ml). After
centrifugation (50,000 × g for 1 h) and dialysis for 4
h against buffer A, the retentate (4.5 ml) is injected
15   onto Mono Q HR 10/10 (Pharmacia) equilibrated with
buffer A. The proteins are eluted with a linear
gradient of 0.4 M NaCl, and the pooled active fractions
are passed through a PD-10 (Pharmacia) column
equilibrated in 30 mM Tris-HCl/5 mM potassium
20   phosphate/5 μM calcium chloride pH 7.6 (buffer B). The
protein solution is fractionated on a Bio-Gel HPHT
(Bio-Rad) column equilibrated in buffer B and eluted
with a gradient of 5 to 350 mM potassium phosphate. The
active fractions are pooled and brought to 500 mM with
25   respect to ammonium sulphate, and then fractionated on
a Phenyl-Superose HR 5/5 (Pharmacia) column eluted with
a decreasing ammonium sulphate gradient. The fraction

containing the activity is finally repurified on a
Mono Q HR 5/5 column in Tris-HCl at pH 7.3. After this
step, the protein is more than 97 % pure. It shows no
contaminant protein in SDS-PAGE. This purity is
5   confirmed by the uniqueness of the NH_2-terminal
sequence. Its molecular weight in this technique is
20,000. The different steps of purification of
cobinamide kinase, with their purification factor and
their yield, are given in Table A.

10   The fractions containing cobinamide kinase
activity also possess cobinamidephosphate
guanylyltransferase activity. Moreover, as shown by the
results presented in the table above, the ratio of
these two activities remains constant in the fractions
15   throughout the purification. Lastly, the purified
protein possesses a very high degree of purity,
exceeding 97 %. These results collectively hence
indicate unambiguously that one and the same protein is
responsible for both successive activities, namely
20   cobinamide kinase and cobinamidephosphate
guanylyltransferase of the pathway of biosynthesis of
cobalamins in Pseudomonas denitrificans.

d) NH_2-terminal sequence of Pseudomonas denitrificans
cobinamide kinase/cobinamidephosphate
25   gunaylyltransferase, and identification of the
Pseudomonas denitrificans structural gene coding for
this activity

The NH$_2$-terminal sequence of <u>Pseudomonas</u>

<u>denitrificans</u> cobinamide kinase/cobinamidephosphate

guanylyltransferase was determined as described above.

Ten residues were identified:

5   Ser-Ser-Leu-Ser-Ala-Gly-Pro-Val-Leu-Val

The NH$_2$-terminal sequence of the COBP protein

(Figure 47) corresponds exactly to this sequence except

that, in the sequence presented in Figure 47, a

methionine precedes the peptide sequence determined by

10   direct sequencing. It follows from this that the amino

terminal methionine is definitely excised in vivo by

methionine aminopeptidase (Ben Bassat and Bauer, 1987).

The molecular weight of the purified cobinamide

kinase/cobinamidephosphate guanylyltransferase is

15   estimated by SDS-PAGE electrophoresis at 20,000. The

COBP protein has a molecular weight deduced from its

sequence of 19,500 (Figure 47). The correspondences

between the NH$_2$-terminal sequences and the molecular

weights indicate clearly that the COBP protein

20   corresponds to cobinamide kinase/cobinamidephosphate

guanylyltransferase. The cobP gene is the cobinamide

kinase/cobinamidephosphate guanylyltransferase

structural gene.

6.2 - Determination of the properties of COB

25   proteins by measurement of accumulated biosynthesis

intermediates

This example illustrates how it is possible to assign an enzymatic activity to a COB protein of Pseudomonas denitrificans. This activity is assigned on the basis of data obtained relating to accumulated biosynthesis intermediates in the Cob mutant or mutants blocked in the step in question. In effect, if a mutant accumulates a biosynthesis intermediate, it is very probable that this mutant is blocked in the step which has the intermediate in question as its substrate.

6.2.1. Properties of the COBC and COBD proteins

The Cob mutants G643 (Agrobacterium tumefaciens) and G572 (Pseudomonas putida) already described in Examples 1 and 4 are blocked in the step corresponding to the COBC protein. In effect, these two mutants are not complemented by the inactivating insertions of transposons Tn5 which occur in the cobC gene. The two strains G643 and G572, as well as the unmutated parent strains [C58-C9 Rif' and KT 2440 Rif' (Cameron et al., 1989)], were cultured in PS4' medium for A. tumefaciens and PS4'' medium for P. putida (PS4' and PS4'' correspond to PS4 medium containing 100-fold and 1000-fold, respectively, less cobalt than PS4 described above) for 3 days as described above. $^{57}CoCl_2$ was added to the cultures (2.5 μCi/0.1 μm for a 25-ml culture). The intracellular corrinoids were isolated in their native form and identified by their HPLC

behaviour. The parent strains do not accumulate

corrinoids other than coenzyme B_{12}. The two mutants G643

and G572 accumulate adenosylated cobyric acid in

respective proportions of 11 % and 6 %. These %

proportions are calculated relative to the level of

coenzyme B_{12} synthesised by the parent strain. Apart

from cobyric acid, mutant G643 accumulates cobyrinic

acid pentaamide in a proportion of 2 %; cobyrinic acid

pentaamide is the intermediate which precedes cobyric

acid. A study of these mutants brings out the fact that

they are blocked after cobyric acid. All these Cob

mutants are blocked either between uro'gen III and

cobinamide, or between cobinamide and the cobalamins.

The mutants G643 and G572 are blocked between uro'gen

III and cobinamide. Now, if these mutants are blocked

before cobinamide, and both accumulate cobyric acid,

the proteins for which they code can participate only

in the enzymatic step (referred to as cobinamide

synthase) which catalyses the amidation of cobyric acid

with an aminopropanol residue to give cobinamide; they

can also possibly participate in the synthesis of the

substrate of the reaction which provides aminopropanol,

if not aminopropanol itself. The cobC gene codes for a

protein which is either cobinamide synthase or one of

its subunits.

The Cob mutant G634 of Agrobacterium

tumefaciens which is blocked in the step corresponding

to the cobD gene was analysed in the same manner. This
mutant is not complemented by the inactivating
insertions in the cobD gene (Example 4.1). The only ·
intracellular corrinoid found in this mutant is
adenosylated cobyric acid. Like the above mutants, this
mutant codes for a protein participating in the
conversion of cobyric acid to cobinamide, or else
possibly in the synthesis of the other substrate of the
reaction.

These two different genes (cobC and cobD)
code for two proteins which participate in the same
step.

6.2.2. Properties of the COBF to COBM
proteins

The Agrobacterium tumefaciens mutants already
described were studied, the study described in Example
4.2 having shown in which genes each of these mutants
is blocked. They are the following mutants: G612
(cobF), G615 (cobG), G616 (cobH), G613 (cobI), G611
(cobJ), G620 (cobK), G638 (cobL) and G609 (cobM); we
have shown in brackets the Pseudomonas denitrificans
gene responsible for the complementation of these
mutants (Example 5), which hence corresponds to the
gene mutated in this mutant. These mutants were
cultured in PS4 medium as described above with labelled
cobalt. After four days' incubation, the mutants were

analysed for their intracellular content of corrinoids
and decobaltocorrinoids (see Examples 6.1.2 and 9).

Table: Intermediates accumulated by Agrobacterium
tumefaciens mutants blocked in the genes of the 8.7-kb
fragment of Pseudomonas denitrificans

Strain	Intracellular decobalto-corrinoids in % [1]			Intracellular corrinoids as % of cobalamins	Mutated gene
	HBA	HBAM	HBAD		
C58-C9*	100	100	100	coenzyme B_{12} 100	–
G612	< 5	< 5	64	cobinamide 2.2 coenzyme B_{12} 34	cobF
G615	< 5	< 5	84	coenzyme B_{12} 17	cobG
G616	35	< 10	< 10	coenzyme B_{12} 13	cobH
G613	< 5	< 5	57	coenzyme B_{12} <1	cobI
G611	< 5	< 5	65	coenzyme B_{12} <1	cobJ
G620	12	< 5	< 10	coenzyme B_{12} <1	cobK
G638	< 5	< 5	47	coenzyme B_{12} <1	cobL
G609	< 5	< 5	33	coenzyme B_{12} <1	cobM

HBA   :   hydrogenobyrinic acid

HBAM :   hydrogenobyrinic acid monoamide

HBAD :   hydrogenobyrinic acid diamide

* in fact, this is strain C58-C9 Rif'Nal' already described
(Cameron et al., 1989)

[1] the values are expressed as % of the same intermediates
accumulated in the unmutated parent strain C58-C9 Rif^{r}Nal^r.

These results show that none of the mutants
accumulate any corrinoid (with the exception of the
mutant inactivated in the cobF gene, G612, which, for

its part, accumulates cobinamide but at a low level

equivalent to 2.2 % of the cobalamins synthesised by

the unmutated strain). However, some mutants (G612,

G615 and G616) have levels of cobalamins which

5   represent more than 10 % of the cobalamin level of the

parent strain. It is probable that all these mutants

are blocked at least before cobyrinic acid diamide. All

the mutants accumulate hydrogenobyrinic acid and

hydrogenobyrinic acid diamide in smaller quantities

10   than the unmutated strain; they are hence very probably

blocked before hydrogenobyrinic acid. It may be

concluded that all the cobF to cobG genes code for

proteins which participate before hydrogenobyrinic

acid. Mutant G613 is known to be mutated in the cobI

15   gene which codes for SP_2MT, participating well before

hydrogenobyrinic acid. For this mutant, the results of

the present example relating to the accumulation of

intermediates are in complete agreement with the step

inactivated in this mutant, namely, this mutant

20   accumulates no intermediate after hydrogenobyrinic acid

at a level higher than that observed with the unmutated

strain. This result is, for the cobF, cobJ, cobL and

cobM genes, consistent with those of Example 6.4, where

it proposed that these genes code for proteins which

25   catalyse SAM-dependent transfers of methyl and hence

which participate before hydrogenobyrinic acid. With

the exception of cobI, which is the SP_2MT structural

gene, these genes participate after precorrin-3. In effect, since they are neither the structural genes for SUMT nor for SP_2MT, they inevitably participate later, that is to say after precorrin-3 (all the cob genes

5   described in the present invention participate between uro'gen III and the cobalamins). These cobF to cobH and cobJ to cobM genes code for enzymes which participate between precorrin-3 and hydrogenobyrinic acid.

       6.2.3. Properties of the COBS and COBT

10   proteins

       The mutant G2035 described in Examples 1 and 4.3 is blocked in the step corresponding to the COBS protein. The mutant G2037 described in Example 1 is blocked in this step corresponding to the COBT protein.

15   These strains, as well as the parent strain (Agrobacterium tumefaciens C58C9Rif'), are cultured in PS4' medium (this is PS4 medium in which the cobalt chloride concentration is 100-fold lower than in PS4 medium) in the presence of radioactive cobalt $^{57}CoCl_2$ for

20   3 days, and their intracellular content of decobaltocorrinoids is analysed, as is the corrinoid content, as already described above (see Example 6.2.2). The strains G2035 and G2037 do not accumulate corrinoids, and large concentrations (greater than

25   those observed with the parent strain) of hydrogenobyrinic acid and hydrogenobyrinic acid mono- and diamide are present only with strain G2035. This

mutant is probably blocked in a step located after
hydrogenobyrinic acid diamide and before cobyrinic acid
diamide. Consequently, the cobS gene is considered to
code for one of the enzymes involved in the conversion
of hydrogenobyrinic acid diamide to cobyrinic acid
diamide; this protein may hence participate either in
the insertion of cobalt, or in the reduction of the
cobalt of unadenosylated cobyrinic acid a,c-diamide. In
contrast, the mutant G2037 is considered to be blocked
in a step located upstream of hydrogenobyrinic acid.
The cobT gene is considered to code for a protein
involved in an enzymatic step upstream of
hydrogenobyrinic acid and downstream of precorrin-3
(other structural genes coding for enzymes involved
downstream of precorrin-3 have already been
identified). Another possibility for the COBT protein
is that it participates, as proposed in Example 5, as a
cobalt-binding protein and/or as a protein which
interacts with other protein(s) via its acidic portion.

6.2.4. Properties of the COBV protein

The mutants G2039 and G2040 described in
Examples 1 and 4.4 are blocked in the step
corresponding to the COBV protein. These strains, as
well as the parent strain, are cultured in PS4' medium
for 3 days in the presence of radioactive cobalt $^{57}CoCl_2$,
and their intracellular content of decobaltocorrinoids
is then analysed and the corrinoid content is

determined as described in Example 9. Strains G2039 and
G2040 accumulate cobyric acid, cobinamide, cobinamide
phosphate and GDP-cobinamide. These mutants are
probably blocked in an enzymatic step downstream of

5   GDP-cobinamide. The cobV gene is considered to code for
an enzyme involved in the conversion of GDP-cobinamide
to cobalamin, see Figure 5. This result is in complete
agreement with the cobalamin-5'-phosphate synthase
activity of the COBV protein which possesses Ado-GDP-

10   cobinamide as a substrate.

6.3 - Determination of the activity of COB
proteins by studies of affinity for SAM

This example illustrates how it is possible,
using COB proteins purified from Pseudomonas

15   denitrificans, to demonstrate in vitro a SAM-binding
activity. If a COB protein possesses such an activity,
it means that this COB protein is a methyltransferase
of the pathway, and that it participates in one of the
transfers of the eight methyl groups which occur

20   between the uro'gen III and cobyrinic acid.

6.3.1. Test of affinity for SAM on a purified
protein

The test is based on the principle according
to which methyl transferases of the pathway of

25   biosynthesis of cobalamins definitely have an SAM-
binding site. This site must be demonstrated by a
higher affinity of SAM than for any protein which does

not specifically bind SAM. After incubation of the
protein under study in the presence of an excess of
radioactive SAM, the latter is separated from the free
SAM by gel permeation chromatography. The radioactivity
5   appearing in the fraction having the molecular weight
of the protein corresponds to the SAM bound during the
incubation. The chromatography is performed at 2°C in
order to limit to the maximum the release of bound SAM
during the separation.

10   The protein (approximately 10 μg) is
incubated for 10 minutes at 30°C in 0.1 M Tris-HCl
pH 7.7 (200 μl) with [methyl-^{3}H]SAM (5 nmol; 1 μCi).
After incubation, a portion (100 μl) of the mixture is
immediately injected onto a TSK-125 (Bio-Rad) column
15   eluted at 1 ml/minute with the 50 mM sodium
sulphate/20 mM sodium dihydrogen phosphate mixture,
pH 6.8, recommended by the distributor of this column.
0.5-ml fractions are collected and subjected to liquid
scintillation counting. The retention times of the
20   protein and the SAM are obtained directly from the
recording of the absorbance of the eluate at 280 nm.

6.3.2. In vitro study of the binding of SAM
to the COBA and COBF proteins of Pseudomonas
denitrificans

25   a) Purification of the COBF and COBA proteins
The COBF protein of Pseudomonas denitrificans
is purified as described below. In a typical

purification experiment, wet cells (5 g) of strain
SC150 Rif' into which plasmid pXL1546 has been
introduced (see Ex. 7.3), obtained after culturing in
PS4 medium, are resuspended in 0.1 M Tris-HCl pH 7.7

5   (30 ml) and sonicated for 15 minutes at 4°C. The crude
extract is then recovered by centrifugation for 1 hour
at 50,000 g, and the supernatant is passed through a
DEAE-Sephadex column (1 ml of gel) to remove the
tetrapyrrole compounds present. Proteins (10 mg;

10   0.7 ml) of this extract are then injected onto a MonoQ
HR 5/5 column equilibrated with the same buffer. The
proteins are eluted with a linear KCl gradient (0 to
0.25 M). The COBF protein is eluted with 0.20 M KCl. It
is diluted twofold with 0.1 M Tris-HCl pH 7.7 and

15   purified a second time on a MonoQ HR 5/5. SDS-PAGE
electrophoresis with visualisation with Coomassie blue
is used to reveal the protein. This technique shows,
moreover, that COBF is approximately 95 % pure after
this purification step. The NH_2-terminal sequence of the

20   purified protein was determined as described above. Two
NH_2-terminal sequences appear at the same time in each
degradation cycle; they are the following sequences, in
the proportions indicated:

Sequence 1 (abundance 34 %)

25 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|----|----|
| Ala | Glu | Ala | Gly | Met | Arg | Lys | Ile | Leu | Ile | Ile |

Sequence 2 (abundance 66 %)

1	2	3	4	5	6	7	8	9	10	11
Met	Arg	Lys	Ile	Leu	Ile	Ile	Gly	Ile	Gly	Ser

Sequence 1 corresponds to the NH$_2$-terminal

sequence of the COBF protein which is given in

5   Figure 16, except that the amino-terminal methionine is

excised according to rules already stated (Hirel et

al., 1989) by methionine aminopeptidase (Ben Bassat and

Bauer, 1989). Sequence 2, present in the larger amount,

corresponds to the same protein but having its

10   translation initiation apparently done not at the

translation initiation ATG codon we had assumed, but at

that located 5 codons downstream on the coding frame

(Figure 16). In effect, the amino acids of this

sequence are exactly those which are found in the

15   sequence of the COBF protein starting from the second

methionine (amino acid No. 6) of this sequence (Figure

16). In this case, the amino-terminal methionine is not

excised, which confirms the rules already stated (Hirel

et al., 1989). In strain SC510 Rif' carrying plasmid

20   pXL1546, there are two translation initiations, on the

one hand that corresponding to the methionine codon

positioned at the correct distance, in our

construction, from the Shine-Dalgarno sequence, and on

the other hand that which is carried out at the second

25   methionine codon occurring in the sequence of the cobF

gene presented in Figure 16. It emerges from this that

the COBF protein proably begins not at the methionine

indicated in Figure 16, but at that occurring 5 amino
acids further on.

At all events, this result shows that the
COBF protein is, indeed, the one expressed, and that

5   the latter is expressed in a form elongated by 4 amino
acids. During purification, both protein forms are
purified. In this example, the mixture of these two
purified proteins is referred to by us as purified COBF
protein.

10   The COBA protein of Pseudomonas denitrificans
is purified as described above (Blanche et al., 1989).

b) Binding of SAM

The binding of SAM to these two proteins is
studied as described above in Example 6.3 a). Bovine

15   serum albumin and the purified COBH protein are used as
negative controls. For the COBA and COBF proteins, a
peak of radioactivity is observed at emergence from the
TSK-125 column at the emergence time of these proteins
(Figure 20). In this test, the COBI protein displays

20   the same property of binding of SAM. In contrast, there
are no such peaks of radioactivity with BSA and the
COBH protein. This test demonstrates the in vitro
binding of SAM to the COBA, COBI and COBF proteins.
These results show that COBA, COBI and COBF are SAM

25   methyltransferases. This result is in complete
agreement with the COBA and COBI activities, since they
are the SUMT and the SP$_2$MT, respectively, of Pseudomonas

denitrificans. The COBF protein is hence probably an
SAM methyltransferase of the pathway of biosynthesis of
cobalamins. This test confirms that COBF is a
methyltransferase.

5   6.4 - Determination of the activity of COB
proteins by sequence homology studies

This example illustrates how it is possible
to find the COB proteins which are SAM methyl-
transferases of the pathway of biosynthesis of
10   cobalamins by comparisons between the sequences of
various COB proteins of Pseudomonas denitrificans.

The COBI and COBA proteins are both SAM
methyltransferases of the biosynthetic pathway. These
two proteins were compared according to the programme
15   of Kanehisa, 1984. This comparison brings out three
regions of strong homology (Figure 21). In each of
these regions, there is more than 45 % strict homology
between the two proteins. Three regions of strong
homology between COBA and CYSG are also presented
20   (Figure 22); they are the same regions of COBA which
display a strong homology with COBI. These regions of
strong homologies between COBA, CYSG and COBI display
homology with other COB proteins. The proteins in
question are COBF, COBJ, COBL and COBM (Figure 23). As
25   regards the region 1, the COBF, COBL and COBM proteins
display significant homologies with respect to all the
Genpro proteins, Genpro being a Genbank (version 59)

protein extraction augmented by putative coding

portions larger than 200 amino acids, according to the

programme of Kanehisa (1984). As regards the region 2,

the COBJ, COBL and COBM proteins display significant

homologies with respect to all the Genpro (version 59)

proteins. As regards the third region of homology,

COBJ, COBL and COBM display significant homologies with

respect to all the Genpro (version 59) proteins. The

sequence comparisons hence enable it to be demonstrated

that four proteins, COBF, COBJ, COBL and COBM, display

significant homologies with the conserved regions of

the sequences of three types of methyltransferases,

COBA, COBI and COBF. The COBG, COBH and COBK proteins

do not display significant homologies with the

conserved regions of the methylases. The COBF protein

displays a significant homology with the other proteins

only in the region 1. These homologies must probably

correspond to the fact that all these proteins are

methyltransferases. This result ties up with the

biological data described for COBF, relating to the

capacity possessed by this protein for binding SAM

in vitro (Example 6.3). These homologies on the one

hand enable it to be confirmed that COF is an SAM

methyltransferase of the pathway of biosynthesis of

cobalamins, and on the other hand demonstrate that

COBJ, COBL and COBM could be SAM methyltransferases of

the pathway of biosynthesis of cobalamins. These

results also show the homology existing between the COB

proteins of P. denitrificans and the isofunctional

proteins of other microorganisms.

EXAMPLE 6 (B) - Purification and cloning of the

Methanobacterium ivanovii SUMT structural gene

This example illustrates how it is possible

to obtain, in other microorganisms, COB enzymes and cob

genes corresponding to those identified in

P. denitrificans.

6(B).1. Purification of Methanobacterium

ivanovii SUMT

This example describes the purification of

Methanobacterium ivanovii SUMT and a study of its

catalytic properties.

Methanobacterium ivanovii strain DSM2611 is

cultured as described (Souillard et al., 1988). Wet

cells (12 g) are obtained. The latter are resuspended

in 0.1 M Tris-HCl buffer pH 7.6 (80 ml) containing 5 mM

DTT and 1 mM EDTA, and sonicated for 1 h 30 min at 4°C

and then centrifuged for 1 h at 50,000 g. Free

tetrapyrrole compounds are then cleared from the

extract by passage through a small DEAE-Sephadex A25

column set up in the same buffer. The proteins

precipitating at between 55 and 75 % ammonium sulphate

saturation are solubilised in a 0.1 M Tris-HCl pH 7.5,

0.5 mM DTT, 1.7 M ammonium sulphate buffer and injected

onto a Phenyl-Superose HR 10/10 (Pharmacia France/SA)

column eluted with a decreasing gradient (1.7 M to 0 M with respect to ammonium sulphate). The active fractions are passed through a Sephadex G-25 column equilibrated with 0.1 M Tris-HCl pH 7.5, 0.5 mM DTT,

5 25 % glycerol buffer (buffer A), then injected onto a Mono Q HR 5/5 (Pharmacia France SA) column equilibrated with buffer A and eluted with a KCl gradient of 0 to 0.3 M; this step is repeated a second time under the same conditions. Gel permeation chromatography of the

10 active fraction of the preceding step on Bio-Sil TSK-250 (BioRad France SA) enables a protein which is homogeneous in SDS-PAGE and in RP-HPLC (C-18 μBondapak) to be obtained. The different steps of purification, with their yield, as well as their purification factor,

15 are described in the table below.

As shown in this table, the total purification factor is more than 4,500. Some properties of the pure enzyme have been studied according to methods already described (Blanche et al., 1989). This

20 enzyme does indeed have SUMT activity, i.e. it does indeed catalyse the SAM-dependant transfer of two methyl groups at C-2 and at C-7 of uro'gen III. The molecular weight of the enzyme estimated by gel permeation is 60,000 +/- 1,500, while by SDS-PAGE it is

25 29,000, which shows clearly that it is a homodimeric enzyme. Under conditions already described (Blanche et al., 1989), the enzyme has a Km for uro'gen III of 52

+/- 8 nM. In addition, this enzyme does not show

inhibition by substrate at concentrations below 20 μM,

whereas Pseudomonas denitrificans SUMT shows an

inhibition by uro'gen III at a concentration above 2 μM

(Blanche et al., 1989).

Table: Purification of M. ivanovii SUMT

Purification step	Vol (ml)	Proteins (mg)	Sp. activity (u/mg of proteins)	Yield	Purific- ation factor[1]
Crude extract	92	731	0.337	-	-
55-75 % AS	7.1	153	1.215	76	3.6
Phenyl-Superose	9.5	8.34	15.35	52	46
Mono Q 5/5	1.0	0.252	422	43	1252
Bio-Sil TSK	1.0	0.061	1537	38	4561

[1]/ calculated from the yield of proteins.

The Vmax of M. ivanovii SUMT was determined.

It is 1537 U/mg of proteins. This value is greater than

that found for P. denitrificans SUMT, already

determined under optimal conditions for the reaction

(taking account of its inhibition by uro'gen III), 489

U/mg of proteins (Blanche et al., 1989).

6(B).2. Cloning of the M. ivanovii SUMT

structural gene in E. coli

6(B).2.1. Cloning of a fragment internal to the M.

ivanovii SUMT structural gene. For this purpose, the

procedure is as follows: 200 picomols of M. ivanovii

SUMT are used for the NH_2-terminal sequencing of the

protein as described above. In addition, a peptide

fragment obtained by tryptic digestion of the protein

is likewise subjected to a sequencing of its

NH$_2$-terminal portion. The sequences obtained are

presented in Figure 48. The sense and antisense

5   oligonucleotides 946, 923 and 947, respectively (see

Figure 48) are synthesised as described above; these

oligonucleotides contain a restriction site at their 5'

end, which is either EcoRI for the sense

oligonucleotides or HindIII for the antisense

10   oligonucleotide. These oligonucleotides are used for an

enzymatic DNA amplification experiment (Saiki et al.,

1988) as shown diagrammatically in Figure 48.B.

        M. ivanovii genomic DNA is prepared in the

following manner: M. ivanovii (DSM 2611) cells (0.4 g)

15   are washed with 0.15 M NaCl solution. The cells are

then incubated in a 25 % sucrose, 50 mM Tris-HCl pH 8,

lysozyme (40 mg) solution (4 ml), and thereafter for 2

to 3 h at 50°C after the addition of proteinase K

(40 mg) and a 0.2 % SDS, 0.1 M EDTA pH 8 solution

20   (5 ml). The DNA is then extracted with

phenol/chloroform (50%/50%) twice and then twice with

chloroform, and thereafter precipitated with

isopropanol and taken up in TNE (10 mM Tris-HCl pH 8, 1

mM EDTA, 100 mM NaCl) (3 ml).

25   Enzymatic amplification of M. ivanovii DNA is

performed according to the protocol of Saiki et al.,

1988, in a volume of 0.1 ml with M. ivanovii genomic

DNA (600 ng), using the primers 946 and 947 (reaction
1) or 923 and 947 (reaction 2). The buffer used for
this reaction is 1 mM MgCl$_2$, 50 mM KCl, 0.001 % gelatin
and each dNTP at a concentration of 0.2 mM; for each
amplification reaction, 10 mg of each oligonucleotide
are used, as well as Taq DNA polymerase (2.5 units)
(Cetus Corporation). Amplification is carried out over
30 cycles in the Perkin-Elmer Cetus DNA Amplication
system; during each cycle, the DNA is denatured for
1 min at 95°C, the oligonucleotide primers are
hybridysed with single-stranded DNA for 2 min at 38°C
and the newly formed strands are polymerised for 3 min
at 72°C. The amplification products are then extracted
with chloroform and thereafter undergo ethanol
precipitation; they can then be visualised after
migration on acrylamide gel, and thereafter be digested
with restriction enzymes such as EcoRI and HindIII.

In the case of reaction 1, two fragments are
observed: at 615 bp as well as at 240 bp. As regards
reaction 2, two fragments are also observed: at 630 and
170 bp. The whole of the product of an enzymatic
amplification reaction between the oligonucleotides
946-947 is separated by migration on acrylamide gel;
the 615-bp fragment is purified as described above.
This fragment is then digested with EcoRi and HindIII
in order to make the ends of the fragment cohesive.
This fragment is then ligated with the DNA of the

replicative form of phage M13mp19. The ligation is

transformed into E. coli TG1. Six recombinant clones

containing a 615-bp insert are analysed by sequencing

with the universal primer-20 (Pharmacia SA, France). As

5   shown in Figure 49, when the single-standed DNA of the

recombinant phages which contain 615-bp insert is

sequenced, there must be observed, downstream of the

EcoRI site, a non-degenerate sequence corresponding to

that of the oligonucleotide 946 followed, in the same

10   frame, by a sequence coding for the amino acids

LITLKAVNVLK?ADVVL (? means that, at this position, the

residue could not be determined); this sequence

corresponds to that which, in the NH$_2$-terminal sequence

of SUMT, follows the amino acids corresponding to the

15   oligonucleotide 946 (see Figure 48). For two clones,

there was actually observed, after the EcoRI site, a

sequence able to code for the NH$_2$-terminal region of

Methanobacterium ivanovii SUMT, this sequence beginning

with the arrangement Pro-Gly-Asp-Pro-Glu-Leu which are

20   the amino acids encoded by a sequence containing the

oligonucleotide 946. This observation shows that these

two recombinant replicative forms contain an insert

which corresponds to a fragment internal to the

Methanobacterium ivanovii SUMT structural gene. The

25   replicative form carrying this fragment internal to the

M. ivanovii structural gene is referred to as pG10.

6(B).2.2. Cloning of the <u>Methanobacterium ivanovii</u> SUMT
structural gene

      <u>Methanobacterium ivanovii</u> genomic DNA is
digested with several restriction enzymes (single or
5   double digestions). After digestion, the fragments are
separated by agarose gel electrophoresis and are then
transferred onto a nylon membrane as described above.
After denaturation of the fragments thus transferred
and prehybridisation, a hybridisation is performed with
10 the replicative form pG10 as a ^{32}P-labelled probe, as
described above. It is thus found that a 3.2-kb
fragment emanating from an <u>EcoRI</u>-<u>Bgl</u>II digestion of
<u>Methanobacterium ivanovii</u> hybridises with the probe
(see Figure 50). Genomic DNA (40 μg) of <u>M. ivanovii</u> are
15 then digested with <u>EcoRI</u> and <u>Bgl</u>II and thereafter
separated by migration on agarose gel. The fragments
having a size of between 3 and 3.5 kb are electroeluted
as described above. The fragments thus purified are
ligated with the vector pBKS+ (Stratagene Cloning
20 Systems, La Jolla) digested with <u>Bam</u>HI-<u>EcoRI</u>. The
ligation is transformed into <u>E. coli</u> DH5α (Gibco BRL).
The transformants are selected on LB medium
supplemented with ampicillin and X-gal. 800 white
colonies are subcultured on filters; after growth and
25 then lysis of the bacteria, a colony hybridisation is
performed according to the technique of Grünstein and
Hogness (1975). The probe used is the replicative form

pG10 labelled with ^{32}P. A single positive clone after this hybridisation test with the probe is found. The plasmid DNA of this clone is referred to as pXL1809 (see Figure 56). A digestion of this DNA with _EcoRI-XbaI_ enables a 3.2-kb insert to be visualised, as expected. Plasmid pXL1809 is sequenced on both strands by the technique of Chen and Seeburg (1985). A sequence of 955 bases is obtained (Figure 51). An analysis of the open reading frames leads us to identify an open reading frame from base 34 (ATG) to base 729 (TGA). This open reading frame codes for a protein whose sequence is presented in Figure 52. This protein has a molecular weight of 24,900 (see Figure 53), which is close to the molecular weight of the protein purified from _M. ivanovii_. The NH$_2$-terminal sequence of this protein is exactly that determined for purified _M. ivanovii_ SUMT (see Figure 48 and Figure 52). These observations establish unambiguously that the cloned and sequenced gene is indeed the _M. ivanovii_ SUMT structural gene. Since this activity is assumed to participate in the biosynthesis of corrinoids in all bacteria, this gene is designated _cor_A gene, and the protein encoded by this same gene CORA protein. The hydrophobicity profile of the CORA protein of _M. ivanovii_, produced from the programme of Hopp and Woods (1981), shows that it is, as expected, a hydrophilic protein, as presented in Figure 54. The CORA protein of

M. ivanovii shows a degree of strict homology of more

than 40 % with respect to COBA of P. denitrificans

(Figure 53). This homology extends over practically the

whole of both proteins, since it relates to residues 3

5   to 227 of CORA of M. ivanovii and residues 17 to 251 of

COBA of P. denitrificans. This homology reflects the

structural homologies existing between two proteins

that catalyse the same reaction. The regions which are

most highly conserved between CORA and COBA of P.

10   denitrificans are the same ones as are conserved

between COBA of P. denitrificans and CYSG of E. coli

(Figure 22).

EXAMPLE 7 - Expression of COB proteins

       7.1 - Expression in Pseudomonas denitrificans

15   This example illustrates that the

amplification of a structural gene for a COB protein of

Pseudomonas denitrificans in Pseudomonas denitrificans

leads to amplification of the activity of the COB

protein.

20   7.1.1 - Expression of the COBA protein

       Plasmid pXL557 corresponds to plasmid pXL59

into which the 2.4-kb BglII-EcoRV fragment (at

positions 80 and 2394, respectively, in the sequence of

Figure 7) of the 5.4-kb fragment has been cloned. This

25   fragment contains the cobA and cobE genes.

       Plasmid pXL545 contains only the cobE gene.

Its construction has been described in Example 4.1.

These two plasmids were introduced by
conjugative transfer into SC510 Rif'. Strains
SC510 Rif', SC510 Rif' pXL59, SC510 Rif' pXL557 and
SC510 Rif' pXL545 were cultured in PS4 medium. At 4
days, culturing was stopped and the SUMT activities
were assayed according to a standard protocol already
described (F. Blanche et al., 1989). The activities are
given below.

Table: SUMT activity of SC510 Rif' and of some of its
derivatives

Strain	SUMT assayed nmol/h/mg of protein
SC510 Rif'	0.05
SC510 Rif' pXL59	0.04
SC510 Rif' pXL557	2.10
SC510 Rif' pXL545	0.05

It emerges clearly from these results that
only plasmid pXL557 brings about a marked increase in
SUMT activity (a factor of 50) in SC510 Rif'. This
increase results from the amplification of cobA and not
of cobE, since plasmid pXL545, which permits the
amplification of only cobE, does not produce an
increase in SUMT activity. This result confirms that

cobA is the structural gene for SUMT of Pseudomonas
denitrificans. This result shows that it is possible to
obtain an amplification of the SUMT activity in
Pseudomonas denitrificans by amplification of the
structural gene for SUMT of Pseudomonas denitrificans.

7.1.2 - Expression of the COBI protein

A fragment originating from the 8.7-kb DNA
fragment containing the structural gene for SP_2MT (cobI)
is cloned into a plasmid having a broad host range in
Gram-negative bacteria, and this plasmid is then
introduced by conjugation into Pseudomonas
denitrificans SC510 Rif'. The S-adenosyl-
L-methionine:precorrin-2 methyltransferase activity of
the strain is then measured relative to that of the
strain carrying the vector.

The 1.9-kb BamHI-BamHI-SstI-SstI fragment
containing the cobH and cobI genes is purified from the
8.7-kb fragment. XbaI and EcoRI linkers are placed at
the BamHI and SstI ends, respectively, after the latter
have been filled in with bacteriophage T4 DNA
polymerase. The fragment is then inserted between the
XbaI and EcoRI sites of the broad host range plasmid
pXL59. It carries kanamycin resistance. The plasmid
thereby obtained is designated pXL1148 (Figure 24).

Separately, a related plasmid was
constructed: the 1.5-kb BamHI-BamHI-SstI fragment
containing only the whole cobH gene and the 5' portion

of the cobI gene was purified from the 8.7-kb fragment.
XbaI and EcoRI linkers were added at the BamHI and SstI
sites, respectively, after the latter had been filled
in or digested with phage T4 DNA polymerase. This

5   fragment was then inserted between the EcoRI and XbaI
sites of pXL59 to give plasmid pXL1149. Plasmids
pXL1148 and pXL1149 differ only in the presence in
pXL1148 of the 0.3-kb SstI-SstI fragment which contains
the 3' end of the cobI gene. pXL1148 possesses the

10   whole structural gene for cobI, in contrast to pXL1149.
Both plasmids contain the cobH gene.

These two plasmids were introduced by
conjugation into SC510 Rif'. Strains SC510 Rif', SC510
Rif' pXL59, SC510 Rif' pXL1148 and SC510 Rif' pXL1149 are

15   cultured in PS4 medium. After 4 days of culture, the
cells are harvested and the SP_2MT activities are assayed
as described in Example 6.1.3 a).

The result of these assays is given below,
with the SP_2MT activities defined as in Example

20   6.1.3 a).

Table: SP$_2$MT activities of various strains derived from

Pseudomonas denitrificans

Strain	SP$_2$MT activity[1] in %
SC510 Rif'	< 5
SC510 Rif' pXL59	< 5
SC510 Rif' pXL1148	75
SC510 Rif' pXL1149	< 5

[1] per 500 μg of crude extract introduced in the test.

The activity is expressed in % as defined in Example 6.1.3 a).

Only plasmid pXL1148 brings about a substantial increase in SP$_2$MT activity. In contrast, plasmid pXL1149 does not give results different from those observed with the controls SC510 Rif' and SC510 Rif' pXL59. pXL1148 is the only plasmid to contain the cobI gene, and it is the only one to amplify SP$_2$MT activity; this result confirms that the structural gene for SP$_2$MT of Pseudomonas denitrificans is the cobI gene. Furthermore, if the total proteins of these different strains are separated by electrophoresis under denaturing conditions (SDS-PAGE with 10 % of acrylamide), the presence of a band which corresponds to a protein having a molecular weight of 25,000 is observed specifically in the case of pXL1148 (Figure

25). The molecular weight of this protein corresponds
to that of the COBI protein. Plasmid pXL1148 enables
overproduction of the COBI protein to be obtained in
Pseudomonas denitrificans.

5   7.1.3 - Expression of COBF

The expression is obtained by positioning the
Ptrp promoter of E. coli and the ribosome-binding site
of the cII gene of bacteriophage lambda upstream of the
cobF gene. The expression thereby obtained is much
10   higher than that observed by simple gene amplification
using the same multicopy plasmid.

The 2-kb EcoRI-BamHI-BamHI fragment of
pXL1496 (Example 7.2.1 below) is purified (Figure 26).
This fragment contains the Ptrp promoter of E. coli and
15   the ribosome-binding site of the cII gene of
bacteriophage lambda upstream of the cobF gene.
Dowstream of the cobF gene, there is the terminator of
the rrnB operon of E. coli. This fragment is cloned at
the EcoRI-BamHI sites of plasmid pKT230 to give pXL1546
20   (Figure 26). pKT230 is a plasmid of the incompatibility
group Q which replicates in almost all Gram-negative
bacteria (Bagdasarian et al., 1981); this plasmid
carries kanamycin resistance. Plasmid pXL1546 and
pKT230 are introduced by conjugation into SC510 Rif'.
25   Strains Sc510 Rif', SC510 Rif' pKT230 and SC510 Rif'
pXL1546 are cultured in PS4 medium as described above.
After four days of culture, the total proteins of the

different strains are analysed in 10 % SDS-PAGE. As

shown in Figure 27, a protein of molecular weight

32,000 which is overexpressed is observed in the

extract of SC510 Rif' pXL1546; this protein comigrates

with the protein which is overexpressed by <u>E. coli</u> B

pXL1496 (Example 7.2.1 below). Furthermore, this

protein is specifically expressed in strain SC510 Rif'

containing pXL1546, where it represents at least 20 %

of the total proteins. In contrast, this protein is not

observed in the total proteins of strains SC510 Rif' and

SC510Rif' pKT230. This overexpressed protein is hence

the COBF protein.

7.1.4 - <u>Expression of COBH</u>

This example describes the amplification of a

DNA fragment of <u>Pseudomonas denitrificans</u> containing

the <u>cobH</u> gene. The protein which is encoded by this

gene is purified; it is the COBH protein. Plasmid

pXL1149, described in Example 7.1.2, contains in the

DNA insert originating from the 8.7-kb fragment only

the whole <u>cobH</u> gene. In SC510 Rif', this plasmid, in

contrast to the vector, brings about the overexpression

of a protein of molecular weight 22,000 (Figure 25).

7.1.5 - <u>Expression of COBV</u>

This example describes the amplification of

cobalamin-5'-phosphate synthase activity by a plasmid

carrying only cobV (pXL699, see Figure 38). The

cobalamin-5'-phosphate synthase activity is amplified

in SC877 Rif' by plasmid pXL699 by a factor of 50 relative to the same strain with the vector pXL435, pXL1303, pXL1324 or pKT230. This plasmid contains in its insert only the whole of cobV plus the 5'-terminal portions of ORF18 and of cobU. In such a strain (SC877Rif' pXL699), the COBV protein is definitely overexpressed; this overexpression is by a factor of 50 relative to the expression of strain SC877Rif'.

7.1.6 - Expression of the CORA protein

The 1.5-kb EcoRI-BamHI-BamHI fragment of pXL1832 (see Example 7.2.4), containing the Ptrp promotor and then the RBS cII of bacteriophage λ, the M. ivanovii SUMT structural gene and the terminator region of the rrnB operon of E. coli, is cloned at the EcoRI-BamHI sites of pKT230 (Bagdasarian et al., 1981). In this manner, plasmid pXL1841 is obtained (see Figure 56). This plasmid is mobilised in P. denitrificans SC510 Rif' as described above. A transconjugant is studied in greater detail. This strain is cultured in PS4 medium, and the SUMT activity of the bacterial extracts is assayed at the same time as that of the control strain SC510 Rif' pXL435 (Cameron et al., 1989). The activities of these strains are presented below.

Strain	SUMT specific activity in pmol/h/mg of proteins
SC510 Rif'pXL435	50-100

SC510 Rif'pXL1841   1700

This result shows clearly that there is
expression of the SUMT activity of M. ivanovii in P.
denitrificans as a result of plasmid pXL1841, since the
SUMT activity of strain SC510 Rif' pXL1841 is markedly
greater than that of SC510 Rif' pXL435.

7.2 - Expression in E. coli

This example illustrates how a COB protein of
Pseudomonas denitrificans can be overproduced in
E. coli.

7.2.1 - Expression of COBF

The 2250-bp EcoRI-XhoI fragment of the 8.7-kb
EcoRI fragment (at the respective positions 0 and 2250
in the sequence presented in Figure 8) was cloned into
phage M13mp19 (Norrander et al., 1983) between the
EcoRI and SalI sites. The plasmid thereby constructed
is designated pXL1405. An NdeI site was introduced by
directed mutagenesis so that the last three bases (ATG)
of this restriction site constitute the translation
initiation site of the cobF gene. This amounts to
modifying the three bases which precede the ATG of the
cobF gene, GAA (the G is at position 733 in the
sequence presented in Figure 8), to CAT. The NdeI-SphI-
SphI fragment (Figure 26) containing the cobF gene is
then purified; this 1.5-kb fragment is then cloned
between the NdeI-SphI sites of plasmid pXL694 (Denèfle
et al., 1987). The plasmid thereby constructed is

designated pXL1496 (Figure 26). Signals for regulation

of genetic expression in E. coli are present in the

120-bp EcoRI-NdeI fragment (which originates from

pXL694) which precedes the cobF gene. These signals

5   consist of the [-40+1] region of the Ptrp promoter of

E. coli, and then of 73bp which contain the ribosome-

binding site of the cII gene of bacteriophage λ

(Denèfle et al., 1987). Downstream of the cobF gene,

there are the terminators of the rrnB operon of E. coli

10   (in the HindIII-BamHI fragment). Plasmid pXL1496 was

introduced by transformation into the E. coli strain

(Monod and Wollman, 1947). Expression of the cobF gene

was studied as already described (Denèfle et al., 1987)

under conditions where the Ptrp promoter is either

15   repressed (in the presence of tryptophan) or not

repressed (absence of tryptophan). The medium in which

the expression was carried out is M9 minimum medium

(Miller, 1972) supplemented with 0.4 % of glucose,

0.4 % of casamino acids, 10 mM thiamine and 40 μg/ml of

20   tryptophan in the case where it is desired to repress

the Ptrp promoter. E. coli strain B pXL1496 was

cultured at 37°C in the medium described above with

ampicillin (100 μg). As shown in Figure 28, the absence

of tryptophan brings about the expression of a protein

25   of molecular weight 32,000. In effect, in the extract

of total proteins of E. coli B pXL1496 analysed in

SDS-PAGE (Figure 28), a protein of molecular weight

32,000 D which represents between 1 and 4 % of the
total proteins is clearly observed. This protein is
present in markedly smaller quantities in the extract
of the total proteins of E. coli B pXL1496 cultured
under the same conditions but in the presence of
tryptophan. The molecular weight of the protein which
is expressed under these conditions is close to the
molecular weight of the COBF protein deduced from the
amino acid sequence of the protein, which is 28,927
(Figure 16). The protein which is thus expressed in
E. coli is the COBF protein.

7.2.2 - Expression of COBT

Overproduction is obtained by fusing the lac
promotor and the first three codons of lacZ of E. coli
to the 5' end of the cob gene.

The EcoRI site located at position 2624 in
the sequence presented in Figure 32 of the 4.8-kb
fragment contains the fourth codon of the cobT gene.
The 3.5-kb EcoRI-XbaI fragment of pXL837 (see Figure
36) is cloned at the EcoRI and XbaI sites of pTZ18R or
pTZ19R (Pharmacia) to generate pXL1874 or pXL1875,
respectively; these two plasmids differ in the
orientation of the truncated cobT gene with respect to
the promoter of the lactose operon of E. coli (Plac).
Plac is upstream of cobT in pXL1874 while the opposite
is true in pXL1875. Cloning of the EcoRI-XbaI fragment
of pXL837 at the EcoRI-XbaI sites of pTZ18R enables a

protein fusion to be carried out between the first 4
amino acids of E. coli β-galactosidase and the cobT
gene from its 4th codon . Expression of this lacZ' 'cobT
gene is under the control of the expression signals of

5   lacZ. Plasmids pXL1874, pXL1875 and pTZ18R are
introduced by transformation into E. coli strain BL21.
Expression of the cobT gene is studied as already
described (Maniatis et al., 1989).

As shown in Figure 42B, a protein of

10   molecular weight 72,000 is expressed only with pXL1874
and represents, in the extract of total proteins of
BL21, pXL1874 analysed in SDS-PAGE, 1 to 4 % of the
total proteins. The molecular weight of the protein
which is expressed under these conditions is close to

15   the molecular weight of the COBT protein deduced from
the amino acid sequence, which is 70;335, in Figure 40.
This experiment shows clearly that, from the EcoRI site
located in the fourth codon of the cobT gene, an open
reading frame compatible with that found for the cobT

20   gene can be expressed.

7.2.3 - Expression of a truncated COBS
protein

A BamHI site is located at the 45th codon of
the COBS gene. The 1.2-kb BamHI-BamHI fragment

25   containing the 3' portion of the cobS gene and
sequences downstream of this gene is excised from
pXL843 and cloned at the BamHI site of plasmid pET-3b

(Rosenberg et al., 1987) to generate pXL1937. The BamHI fragment is oriented in such a way that the truncated portion of the cobS gene is fused, in frame, with the first 12 codons of the major capsid protein of

5   bacteriophage T7 or gene 10 (Rosenberg et al., 1987). This hydbrid gene is under the control of the ϕ10 promotor of bacteriophage T7. Plasmid pXL1937 and also pET-3b are introduced by transformation into E. coli BL21 pLysS (W. Studier, personal communication). After

10   reisolation on selective medium, both strains are cultured in L liquid medium to an OD at 610 nm of 1; at this stage, the medium is adjusted to an IPTG (isopropyl β-thiogalactoside) concentration of 1mM in order to induce expression of the polymerase of

15   bacteriophage T7 (Rosenberg et al., 1987). The culture is then incubated for 3 h at 37°C and bacterial lysates are thereafter prepared. The total proteins of the bacteria thus cultured are separated by PAGE under denaturing conditions. As seen in Figure 42A, there is

20   specifically overexpression of a 33,000 protein with the culture BL21 pLysS pXL1937. This molecular weight is entirely compatible with the expected molecular weight for the fusion protein (33 kD). This experiment shows clearly that, from the BamHI site located at the

25   45th codon of the cobS gene, an open reading frame compatible with that found for the cobS gene can be overexpressed.

7.2.4. Expression of the CORA protein

The following oligonucleotides were
synthesised as described above:

oligonucleotide 1277

5' GGC CGA ATT CAT ATG GTA GTT TAT TTA 3'

------- 1   2   3   4   5   (1 to 5 first 5

EcoRI _____   codons of M. ivanovii

NdeI   SUMT)

oligonucleotide 1278

5' GGC CGA GCT CTA TTA CAT AAT T

_____===============

SstI

(= sequence appearing in Figure 51, positions 926 to
915, in the strand complementary to the coding strand)

Oligonucleotide 1277 possesses the
recognition sequences for the restriction enzymes EcoRI
and NdeI. The last three bases of the NdeI site (ATG),
which corresponds to a translation initiation codon,
are directly followed by codons 2 to 5 of the M.
ivanovii SUMT structural gene as appear in the sequence
presented in Figure 52. The oligonucleotide 1278
contains the recognition sequence for SstI, followed
directly by the sequence TATTACATAATT which corresponds
to a sequence present in the 955-bp fragment containing
the corA gene presented in Figure 51; this sequence
occurs at position 926 to 915 (see Figure 51) in the
strand complementary to the strand coding of the CORA

protein. The two oligonucleotides 1277 and 1278 hence contain sequences in their 3' portion corresponding, respectively, to the coding strand of the corA gene and to the complementary strand downstream of this gene. These two oligonucleotides may be used to carry out an enzymatic amplification experiment with plasmid pXL1809 as template. This experiment makes it possible to obtain a 910-bp fragment containing the corA gene of M. ivanovii possessing an NdeI site at the ATG of the corA gene, and an SstI site at the other end of the fragment after the end of the corA gene. Enzymatic amplification is carried out as described above for the enzymatic amplification performed on the genomic DNA of M. ivanovii, except that the template consists of DNA (10 ng) of plasmid pXL1809; the temperatures used are the same, but only 20 amplification cycles are carried out. As described above, the amplification products are digested with NdeI and SstI before being separated by migration on agarose gel. As expected, a fragment 910 bp in size is indeed visualised. This fragment is purified as already described. This fragment is cloned at the NdeI and SstI sites of pXL694 (Denèfle et al., 1987). The resulting plasmid, designated pXL1832, is described in Figure 56. In this plasmid, in the same way as described in Example 7.2, the M. ivanovii SUMT structural gene is preceded by the ribosome binding site of the cII gene of bacteriophage λ. Upstream of

this RBS there is the Ptrp promotor. Plasmid pXL832 is introduced into E. coli B5548, which is an E. coli strain carrying the mutation cysG44 (Cossart and Sanzey, 1982) by transformation. The SUMT activities of the strains E. coli B5548 pUC13 and E. coli B5548 pXL1832 are assayed on extracts obtained from cells cultured in LB medium supplemented with ampicillin. The assay of SUMT activity is carried out as already described (Blanche et al., 1989). The results of this assay are given below.

Strain	SUMT specific activity in pmol/h/mg of proteins
E. coli B5548 pUC13	5.9
E. coli B5548 pXL1832	310

The results presented in the table above show clearly that there is expression of a SUMT activity in E. coli strain B5548 when the latter contains a plasmid pXL1832 which expresses M. ivanovii SUMT. The M. ivanovii SUMT can hence be expressed in E. coli.

EXAMPLE 8 - Amplification of the production of cobalamins by recombinant DNA techniques

8.1 - Amplification in P. denitrificans

This example illustrates how an improvement in the production of cobalamins is obtained in Pseudomonas denitrificans SC510 Rif' by amplification of cob genes of Pseudomonas denitrificans SC510.

8.1.1 Improvement in the production of
cobalamins in Pseudomonas denitrificans by removal of a
limiting step in the biosynthesis of cobalamins

This example illustrates how the productivity
of cobalamins in Pseudomonas denitrificans strains may
be improved by amplification of cob genes of
Pseudomonas denitrificans. This improvement results
from the removal of a limiting step of the biosynthetic
pathway.

Plasmid pXL367 is described in Example 4.2
(Figure 13). This plasmid corresponds to pRK290 (Ditta
et al., 1981) into which the 8.7-kb EcoRI fragment has
been inserted. This plasmid pXL367 effects an
improvement in the biosynthesis of cobalamins in strain
SC510 Rif'. Strains SC510 Rif', SC510 Rif' pRK290 and
SC510 Rif' pXL367 are cultured in an Erlenmeyer in PS4
medium according to the conditions described in the
experimemntal protocols. An improvement in the
production titre due to the presence of plasmid pXL367
is observed. In effect, strain SC510 Rif' pXL367
produces 30 % more cobalamins than strains SC510 Rif'
and SC510 Rif' pRK290. This improvement is not due to
the amplification of unspecified genes of Pseudomonas
denitrificans, but to the specific amplification of the
genes carried by the 8.7-kb EcoRI fragment. In effect,
plasmid pXL723 described in Figure 11 gives no
improvement, and the same production titre is observed

with this plasmid as with strains SC510 Rif' and SC510
Rif' pRK290.

8.1.2 Improvement in the production of
coenzyme B_{12} in Pseudomonas denitrificans by removal of

5   two limiting steps in the biosynthesis of cobalamins

This example illustrates how the productivity
of cobalamins in strains of Pseudomonas denitrificans
may be improved by amplification of cob genes of
Pseudomonas denitrificans. This improvement results

10   from the removal of two limiting steps of the
biosynthetic pathway.

The 2.4-kb ClaI-Eco RV fragment derived from
the 5.4-kb fragment (containing the cobA and cobE
genes) is cocloned with the 8.7-kb EcoRI fragment into

15   the broad host range plasmid pXL203. The plasmid
thereby constructed is referred to as pXL525 (Figure
29). This plasmid is introduced into SC510 Rif' by
conjugation. Strain SC510 Rif' pXL525 produces 20 % more
cobalamins than SC510 Rif' pXL367. Amplification of the

20   cobA and cobE genes enables a further limiting step in
SC510 Rif' in the biosynthesis of cobalamins to be
removed. Pseudomonas denitrificans strain SC510 Rif' is
improved in the present example by the successive
removal of two limiting steps. This example shows that

25   the removal of two limiting steps in the biosynthesis
of cobalamins can lead to further improvements in
production.

8.2 - Improvement in the productivity of cobalamins in Agrobacterium tumefaciens

This example illustrates the improvement in the production of cobalamins in a strain productive of

5   cobalamins by amplification of the cob genes of Pseudomonas denitrificans SC510.

The strain used is a strain of a Gram-negative bacterium; it is a strain of Agrobacterium tumefaciens.

10   The plasmids described in Examples 4.2 and 8.1, pXL367 and pXL525, as well as the vector pRK290 (Ditta et al., 1981) and plasmid pXL368 (Figure 29), are introduced by conjugative transfer into Agrobacterium tumefaciens strain C58-C9 Rif' (Cameron et

15   al., 1989). Strains C58-C9 Rif', C58-C9 Rif' pRK290, C58-C9 Rif' pXL367, C58-C9 Rif' pXL368 and C58-C9 Rif' pXL525 are cultured in PS4 medium at 30°C as described above. The cobalamins produced are assayed as described above. The production titres are given in the table

20   below.

Table: Titres of vitamin B_{12} produced by different recombinant strains of Agrobacterium tumefaciens

Strain	Vitamin B_{12} in mg/l
C58-C9 Rif'	0.4
C58-C9 Rif' pRK290	0.4
C58-C9 Rif' pXL367	0.8
C58-C9 Rif' pXL368	0.8
C58-C9 Rif' pXL525	1.2

As is clearly apparent in the above table, the production of cobalamins is improved in the Agrobacterium tumefaciens strain used. Two different plasmids improve the production of cobalamins in the Agrobacterium tumefaciens strain used: pXL367 and pXL368. These plasmids contain the 8.7-kb EcoRI fragment (cobF to cobM genes) and the 2.4-kb ClaI-EcoRV fragment (cobE and cobA gene), respectively. Separately, they improve the production of cobalamins by Agrobacterium tumefaciens C58-C9 Rif' by a factor of 2; this result shows that it is possible to improve the production of cobalamins by a strain of Agrobacterium tumefaciens by amplifying fragments carrying cob genes of Pseudomonas denitrificans. In the present case, it is possible to speak of heterologous improvment, that is to say improvement of the production of cobalamins

by one strain by means of the amplification of cob
genes of another strain.

The improvements in production of cobalamins
provided by the different Pseudomonas denitrificans
fragments containing cob genes are capable of
cumulation, i.e., by putting into the same plasmid the
two fragments which are separately cloned into pXL367
and pXL368, additive improvements are observed, in the
sense that plasmid pXL525 provides in Agrobacterium
tumefaciens C58-C9 Rif' an improvement in the production
greater than that provided by each of the fragments
cloned separately into the same vector.

8.3 - Improvement in the productivity of
cobalamins in Rhizobium meliloti

This example describes the improvement in the
production of cobalamins by another strain productive
of cobalamins.

The plasmid described in Example 8.2, pXL368,
as well as the vector pRK290 (Ditta et al., 1981), are
introduced by conjugative transfer into Rhizobium
meliloti strain 102F34 Rif' (Leong et al., 1982). The
transconjugants, namely 102F34 Rif', 102F34 Rif' pRK290
and 102F34 Rif' pXL368, are cultured in PS4 medium at
30°C as described above. The cobalamins produced are
assayed as described above. The production titres are
given in the table below.

Table: Titres of cobalamins produced by different
recombinant strains of Rhizobium meliloti

Strain	Vitamin B_{12} in mg/l
102F34 Rif'	0.4
102F34 Rif' pRK290	0.4
102F34 Rif' pXL368	0.8

As is clearly apparent in the above table,
the production of cobalamins is improved in the
Rhizobium meliloti strain used. Plasmid pXL368 improves
the production of cobalamins by the Rhizobium meliloti
strain used. This plasmid contains the 2.4-kb
ClaI-EcoRV fragment (cobA and cobE genes); it improves
the production of cobalamins by Rhizobium meliloti
102F34 Rif' by a factor of 2. This result shows that it
is possible to improve the production of cobalamins by
a strain of Rhizobium meliloti by amplifying fragments
carrying cob genes of Pseudomonas denitrificans. In the
present case, it is possible to speak of heterologous
improvement, that is to say improvement of the
production of cobalamins by one strain by means of the
amplification of cob genes of another strain.

EXAMPLE 9 - Assay of corrinoids and decobaltocorrinoids
in musts and cells of strains productive of corrinoids

This example illustrates how it is possible
to identify and assay the different corrinoids and
decobaltocorrinoids produced by different strains

productive of cobalamins. This assay makes it possible, inter alia, to assay coenzyme B_{12}.

The musts (or the cells alone) are cyanide-treated as already described (Renz, 1971). After

5 centrifugation, an aliquot of the supernatant is passed through a DEAE-Sephadex column which is then washed with 0.1 M phosphate buffer. The collected fractions are combined and desalted on a Sep-Pak C-18 (Waters) cartridge. After evaporation and resuspension in water

10 (100 μl to 1 ml depending on the quantity of corrinoids present), the corrinoids are identified and assayed by HPLC on a Nucleosil C-18 column (Macherey-Nagel). The column is eluted at 1 ml/min with an acetonitrile gradient (from 0 % to 100 %) in 0.1 M potassium

15 phosphate buffer containing 10 mM KCN.

The corrinoids are visualised by UV detection at 371 nm and/or by specific detection of ^{57}Co (if culturing has been performed in the presence of $^{57}CoCl_2$) using a Berthold LB 505 detector. They are hence

20 identified by comparison of their retention times with standards. Similarly, the "metal-free corrinoids" (hydrogenobyrinic acid, hydrogenobyrinic acid monoamide and hydrogenobyrinic acid diamide) are visualised by UV detection at 330 nm. By this technique, the following

25 intermediates are separated: cobyrinic acid, cobyrinic acid monoamide, cobyrinic acid diamide, cobyrinic acid triamide, cobyrinic acid tetraamide, cobyrinic acid

pentaamide, cobyric acid, cobinamide, cobinamide phosphate, GDP-cobinamide, B_{12} phosphate and vitamin B_{12}. The adenosylated forms of these products are also separated and assayed by this technique. For this purpose, the initial step of the cyanide treatment is cut out and the HPLC column is eluted with buffer devoid of KCN. Figure 31 gives the retention times of different standards separated by this system and identified at emergence from the column by UV absorbance.

A sample of strain SC510 Rif' was deposited on 30th January 1990 at the Centraal Bureau voor Schimmelcultures at Baarn (Netherlands), where it was registered under reference CBS 103.90.